

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 10:42:18 ; Search time 342.5 seconds
(without alignments)

11066.024 Million cell updates/sec

Title: US-09-912-559-1

Perfect score: 1683

Sequence: 1 atgtttccagatgtctga.....aaagtgaagtgtctctaa 1683

Scoring table: IDENTITY_NUC

Gapop 10-0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.GeneSeq_101002:.*
1: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1980.DAT:.*
2: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1981.DAT:.*
3: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1982.DAT:.*
4: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1983.DAT:.*
5: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1984.DAT:.*
6: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1985.DAT:.*
7: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1986.DAT:.*
8: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1987.DAT:.*
9: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1988.DAT:.*
10: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1989.DAT:.*
11: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1990.DAT:.*
12: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1991.DAT:.*
13: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1992.DAT:.*
14: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1993.DAT:.*
15: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1994.DAT:.*
16: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1995.DAT:.*
17: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1996.DAT:.*
18: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1997.DAT:.*
19: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1998.DAT:.*
20: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA1999.DAT:.*
21: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA2000.DAT:.*
22: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA2001A.DAT:.*
23: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA2001B.DAT:.*
24: /SID52/gcgdata/geneSeq/geneSeqn-emb1/NA2002.DAT:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1683	100.0	1683	24	AA145696
2	1683	100.0	3008	24	ABN95170
3	1679.8	99.8	1683	24	AA145697
4	1678.2	99.7	3623	21	AACT6693
5	526.4	31.3	617	21	AA44763
6	242.6	14.4	397	24	AB199281
7	120.2	7.1	451	22	ABAS8323
8	120.2	7.1	451	22	AAK06420
9	120.2	7.1	22	AAK32096	Human bone marrow

c	10	120.2	7.1	451	22	AA137947	Probe #6633 used t
c	11	120.2	7.1	451	24	ABSO6863	Human genome-deriv
c	12	117.6	7.0	2033	15	AAO63951	Hepatocyte growth
c	13	117.6	7.0	2036	24	ABN95114	Gene #1612 used to
c	14	117.6	7.0	2036	24	AB165242	Lung cancer relate
c	15	117	7.0	117	22	ABA70919	Human foetal liver
c	16	117	7.0	117	22	AAK19188	Human brain expres
c	17	117	7.0	117	22	AAK45153	Human bone marrow
c	18	117	7.0	117	22	AA151106	Probe #19792 used
c	19	117	7.0	117	24	ABSI9403	Human genome-deriv
c	20	96.2	5.7	970	15	AAO63945	Gene comprising HG
c	21	90.6	5.4	2512	12	AAO12867	JM1-229 cell line
c	22	87.6	5.2	2252	11	AAO05673	v-PA.alpha1. Desm
c	23	86.8	5.2	1323	8	AAAT0992	Modified tissue pl
c	24	85.2	5.1	1422	8	AAAT0991	v-PA.alpha1. Desm
c	25	84.8	5.0	2137	11	AAO05675	v-PA.alpha2. Desm
c	26	79.6	4.7	2257	11	AAO05674	Chicken urokinase
c	27	78.8	4.7	919	18	AAAT90048	Mutated CDNA codin
c	28	78.6	4.7	1689	10	AAAN91217	Vampire bat plasmi
c	29	78	4.6	1620	11	AAO00543	T-PA with -ve char
c	30	77	4.6	2097	12	AAO12071	T-PA variant havin
c	31	77	4.6	2100	12	AAO12074	tPA-2 CDNA. Synth
c	32	75.4	4.5	1047	14	AAO40658	mt-PA6-E. Homo sa
c	33	75.4	4.5	1065	18	AAAT87456	DNA encoding t-PA
c	34	75.4	4.5	1065	18	AAAT87456	Human truncated tP
c	35	75.4	4.5	1068	14	AAO53320	tPA-1 CDNA. Synth
c	36	75.4	4.5	1068	14	AAO53322	tPA-6 CDNA. Synth
c	37	75.4	4.5	1068	14	AAO40657	tPA-11 CDNA. Synth
c	38	75.4	4.5	1068	14	AAO40660	tPA-12 CDNA. Synth
c	39	75.4	4.5	1068	14	AAO40667	tPA-13 CDNA. Synth
c	40	75.4	4.5	1068	14	AAO40671	tPA-14 CDNA. Synth
c	41	75.4	4.5	1068	14	AAO40673	tPA-16 CDNA. Synth
c	42	75.4	4.5	1068	14	AAO40677	tPA-17 CDNA. Synth
c	43	75.4	4.5	1068	14	AAO40679	tPA-18 CDNA. Synth
c	44	75.4	4.5	1068	14	AAO40681	
c	45	75.4	4.5	1068	14	AAO40681	

ALIGNMENTS

RESULT 1	AA145696	standard: DNA: 1683 BP.
ID	AA145696	
XX	AA145696	
AC	AA145696	
XX	AA145696	
DT	13-JUN-2002	(first entry)
XX	XX	
DE	Human blood coagulation factor VII activating protease DNA.	
XX	XX	
KW	Human: blood coagulation factor VII activating protease: FSAP.	
KW	single-chain plasminogen activator; bleeding disorder; haematological;	
KW	haemostatic; gene: ds.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
FT	Key	Location/Qualifiers
FT	CDS	1..1683
FT	FT	/*tag= a
FT	FT	/product= "FSAP"
XX	XX	
PN	EP1182258-A1.	
XX	XX	
PD	27-FEB-2002.	
XX	XX	
PF	05-JUL-2001: 2001EP-0115691.	
XX	XX	
PR	26-JUL-2000: 2000DE-1036641.	
PR	10-OCT-2000: 2000DE-1050040.	
PR	21-OCT-2000: 2000DE-1052319.	
PR	12-APR-2001: 2001DE-1018706.	
XX	XX	

ID	ABN95170	standard; DNA; 3008 BP.
XX	ABN95170;	
XX	13-AUG-2002 (first entry)	
XX	Gene #1668 used to diagnose liver cancer.	
XX	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;	
XX	metastatic liver tumour; cytosolic; expression profile; disease state;	
XX	disease progression; drug toxicity; drug efficacy; drug metabolism.	
XX	Homo sapiens.	
XX	NC0200229103-A2.	
XX	11-APR-2002.	
XX	02-OCT-2001; 2001MO-US30589.	
XX	02-OCT-2000; 2000US-237054P.	
XX	(GENE-) GENE LOGIC INC.	
XX	Horne D, Alvares C, Peres-da-Silva S, Vockley JG;	
XX	WPI; 2002-426119/45.	
XX	Diagnosing and detecting the progression of liver cancer,	
XX	hepatocellular carcinoma or metastatic liver tumor in a patient,	
XX	involves detecting the level of expression of two or more genes in a	
XX	liver tissue sample	
XX	Claim 1; SEQ ID NO 1668; 298bp; English.	
XX	The invention relates to a novel method for diagnosing and detecting the	
XX	progression of liver cancer, hepatocellular carcinoma or metastatic liver	
XX	tumour in a patient, and differentiating metastatic liver cancer from	
XX	hepatocellular carcinoma in a patient, involving detecting the level of	
XX	expression of two or more genes represented in ABN93503-ABN97455 in a	
XX	tissue sample. The method of the invention has hepatotropic, and	
XX	cytosolic activity. The method is useful for diagnosing and detecting	
XX	the progression of liver cancer, hepatocellular carcinoma and metastatic	
XX	liver carcinoma in a patient. The method is useful for identifying	
XX	expression profiles which serve as useful diagnostic markers as well as	
XX	markers that can be used to monitor disease states, disease progression,	
XX	drug toxicity, drug efficacy and drug metabolism.	
XX	Note: The sequence data for this patent did not form part of the printed	
XX	specification. The sequence data for this patent did not form part of the printed	
XX	at ftp://wipo.int/pub/published_pct_sequences.	
XX	Sequence 3008 BP; 791 A; 837 C; 703 G; 677 T; 0 other:	
XX	Query Match 100.0%; Score 1683; DB 24; Length 3008;	
XX	Best Local Similarity 100.0%; Pred. NO. 0;	
XX	Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
0Y	1 ATGTTTGCCAGGATGTCTGATCTCCATGTTGCTGTTAATGCTCTGCTGGAAAGACA 60	
Db	97 ATCTTTGCCAGGATGTCTGATCTCCATGTTGCTGTTAATGCTCTGCTGGAAAGACA 156	
0Y	61 GCGTGGGGTTCCTCCGATGATCTTTAATGGAAAGCCGTGACACAGCTGACCCCTGAC 120	
Db	157 GCGGTGGGGTTCCTCCGATGATCTTTAATGGAAAGCCGTGACACAGCTGACCCCTGAC 216	
0Y	121 CAGTAATGATTAAAGTACAGAGATTTAATTCAGGAAGAAACACAGTGCACACTTAC 180	
Db	217 CAGTAATGATTAAAGTACAGAGATTTAATTCAGGAAGAAACACAGTGCACACTTAC 276	
0Y	181 CATGCTGAGATCTGACTGAGTGAATACAGTGAAGAACCAAGCTGATTCACGACCCAC 240	

Db	277	CATGCTGACAAATTCCTGACACTGGTACTACACTGAGGACCAAGCTGATCCATGCGACGCCAAC	336
Qy	241	CCCTGTGAAACACGGTGGGAGACTGCTCTGTGCATGGGAGACATTTACATGTACGTGCTG	300
Db	337	CCCTGTGAAACACGGTGGGAGACTGCTGTGTGCATGGGAGACACTTTCATGTACGTGCTG	396
Qy	301	GCCTCTTTCTCTGGGAAATAGTGTCCGAAAGTGCAGAAATACGTGTCAAGGACCAACCATGT	360
Db	397	GTCTCTTTCTCTGGGAAATAGTGTCCGAAAGTGCAGAAATACGTGTCAAGGACCAACCATGT	456
Qy	361	GGCCGGGGCCATGTCTCATATTACCCAGAGATCCCTGCTCTATACGGCGTGTGCTGTAAAC	420
Db	457	GGCCGGGGCCATGTCTCATATTACCCAGAGATCCCTGCTCTATACGGCGTGTGCTGTAAAC	516
Qy	421	CCCTTACACAGTGTCCAGCTGCTGCCAAGTGTCTCTGTATGCAGGCCAAACCCCTGGCAG	480
Db	517	CCCTTACACAGTGTCCAGCTGCTGCCAAGTGTCTCTGTATGCAGGCCAAACCCCTGGCAG	576
Qy	481	AATGGGGCTACTGCTGTCCCGGACATAGCGGAGATCCAGTTCACCTGTGCTGTCCGAC	540
Db	577	AATGGGGCTACTGCTGTCCCGGACATAGCGGAGATCCAGTTCACCTGTGCTGTCCGAC	636
Qy	541	CAGTTCACAGGGAATTTCTGTGAATATGGTCTGTATGTACTGTATGTGGCATGGCTAC	600
Db	637	CAGTTCACAGGGAATTTCTGTGAATATGGTCTGTATGTACTGTATGTGGCATGGCTAC	696
Qy	601	TCCTTACCAGGGAATAATAGACAGTCAACGACGATGCGTGTCTTACTGTGACTTC	660
Db	697	TCCTTACCAGGGAATAATAGACAGTCAACGACGATGCGTGTCTTACTGTGACTTC	756
Qy	661	CACCTCCCTCTTCCAGAGAAATTTACCATGTATTATGAGGAGTGTGTAAACCATGGAGTT	720
Db	757	CACCTCCCTCTTCCAGAGAAATTTACCATGTATTATGAGGAGTGTGTAAACCATGGAGTT	816
Qy	721	GGGGAACACAAATTTCTGTGAGAAACCCAGATGGGAGCAAAAGCCCTGTGTATTATAA	780
Db	817	GGGGAACACAAATTTCTGTGAGAAACCCAGATGGGAGCAAAAGCCCTGTGTATTATAA	876
Qy	781	GTTACCAATGTACAGAGTCAAAATGTGGCAATACTGTGATGTCTAGCCTCTCAGCCAGGAC	840
Db	877	GTTACCAATGTACAGAGTCAAAATGTGGCAATACTGTGATGTCTAGCCTCTCAGCCAGGAC	936
Qy	841	GTTGCCCTACCCAGAGGAAGCCCACTGAGACCATCAACCAAGCTTCGGGGGTTTGACTCC	900
Db	937	GTTGCCCTACCCAGAGGAAGCCCACTGAGACCATCAACCAAGCTTCGGGGGTTTGACTCC	996
Qy	901	TGTGTGAAGACTGTGAATGTGCAGAGAGAGAGATCAAGAAGATCTATGAGGCTTTAAGAC	960
Db	997	TGTGTGAAGACTGTGAATGTGCAGAGAGAGAGATCAAGAAGATCTATGAGGCTTTAAGAC	1056
Qy	961	ACGGCGGCAAGACCCCATGTGCAGAGGCTCCCTCCAGTCCCTGCGCTGTGACCATCTCC	1020
Db	1057	ACGGCGGCAAGACCCCATGTGCAGAGGCTCCCTCCAGTCCCTGCGCTGTGACCATCTCC	1116
Qy	1021	ATGTCCCGAGGGCACTTCTGTGTGTGTGGGGCCGTATGCACCCCTGCTGGGTGTCACTGCT	1080
Db	1117	ATGTCCCGAGGGCACTTCTGTGTGTGTGGGGCCGTATGCACCCCTGCTGGGTGTCACTGCT	1176
Qy	1081	GCCCATGTGCACCGACATATAAAACCAAGCATCTTAAAGGTGTGTGTAGGGGACAGGACTG	1140
Db	1177	GCCCATGTGCACCGACATATAAAACCAAGCATCTTAAAGGTGTGTGTAGGGGACAGGACTG	1236
Qy	1141	AAGAAAGACAAATTTCAATGTACACAGCTTTTATGGGTGTGAGAGATATTCAAGTCAAGCCAC	1200
Db	1237	AAGAAAGACAAATTTCAATGTACACAGCTTTTATGGGTGTGAGAGATATTCAAGTCAAGCCAC	1296
Qy	1201	TACAAATGAAGAGATGTAGCTTCCCAACAATGAATATGTGATGTGTCAAGTTAAAGCCAGTG	1260
Db	1297	TACAAATGAAGAGATGTAGCTTCCCAACAATGAATATGTGATGTGTCAAGTTAAAGCCAGTG	1356
Qy	1261	GATGTCACTGTGTCTGTAGAAATCCAAATACGTGTGAAAGCTGTGTGCTTGTGCTGATGGTCC	1320

QY	781	GTACCAATACACAAAGGTAAATGGGAATACTGTGATCTCTACACCTGGTACGCCACAGAC	840
Db	781	GTTACCAATACACAAAGGTAAATGGGAATACTGTGATCTCTACACCTGGTACGCCACAGAC	840
QY	841	GTTCGCTACCCACAGAGAGAAAGCCCCACTGAGCCATCACAACAGCTTCGGGGTTTGACTCC	900
Db	841	GTTCGCTACCCACAGAGAGAAAGCCCCACTGAGCCATCACAACAGCTTCGGGGTTTGACTCC	900
QY	901	TGTGAAAGACTAGATAGCGAGAGAGAGATCAAGAGAAATCTATGAGAGGCTTTTAAAGAC	960
Db	901	TGTGAAAGACTAGATAGCGAGAGAGAGATCAAGAGAAATCTATGAGAGGCTTTTAAAGAC	960
QY	961	ACGGGGGACAAAGACCCATGACAGGGGTCCCTCCAGTCTCTGCTGCTCTACACATCCG	1020
Db	961	ACGGGGGACAAAGACCCATGACAGGGGTCCCTCCAGTCTCTGCTGCTCTACACATCCG	1020
QY	1021	ATGCCCCAGGGGACACTCTGTGTGTGGGGGGCGGTGATCCACCCCTGCTGGGTCTACTGCT	1080
Db	1021	ATGCCCCAGGGGACACTCTGTGTGTGGGGGGCGGTGATCCACCCCTGCTGGGTCTACTGCT	1080
QY	1081	GGCCACTGCACGACACATTTAAAAACCAACATCTTAAAGGTGTGCTAAGGGACACAGACTTG	1140
Db	1081	GGCCACTGCACGACACATTTAAAAACCAACATCTTAAAGGTGTGCTAAGGGACACAGACTTG	1140
QY	1141	AAGAAACAAATATTCATGACACAGACTTTAGGGTGGACAGAAATATTCAGTACAGCAC	1200
Db	1141	AAGAAACAAATATTCATGACACAGACTTTAGGGTGGACAGAAATATTCAGTACAGCAC	1200
QY	1201	TGCATGTGAAGAGATAGATGATTCGCCCAATGATATGATGCTCAAGTTTAAACCCAGTG	1260
Db	1201	TGCATGTGAAGAGATAGATGATTCGCCCAATGATATGATGCTCAAGTTTAAACCCAGTG	1260
QY	1261	GATGTCACACTGTCTCTAGAAATCCAAATACGTAAGTGTGCTGTGCTGATGGGTCC	1320
Db	1261	GATGTCACACTGTCTCTAGAAATCCAAATACGTAAGTGTGCTGTGCTGATGGGTCC	1320
QY	1321	TTTCCCTCTGGGAGTGAAGTGCACACATCTGTGGCTGGGGTGTACAGAAACAGAAAGGG	1380
Db	1321	TTTCCCTCTGGGAGTGAAGTGCACACATCTGTGGCTGGGGTGTACAGAAACAGAAAGGG	1380
QY	1381	TCCCGCACAGTCTGTGATGATGCCAAAGTCAAGCTGATTTGCCAAGCTTGTGCACATCCCG	1440
Db	1381	TCCCGCACAGTCTGTGATGATGCCAAAGTCAAGCTGATTTGCCAAGCTTGTGCACATCCCG	1440
QY	1441	CACGCTATATACACATGATTTGATGACAGATATGATCTGTGCAGAAATCTTCACAAACCT	1500
Db	1441	CACGCTATATACACATGATTTGATGACAGATATGATCTGTGCAGAAATCTTCACAAACCT	1500
QY	1501	GGGCAACACACTGCGCAGGCTGACTCTGGAGGCCCTTGACTCTGTGAGAGAGAGCGCAC	1560
Db	1501	GGGCAACACACTGCGCAGGCTGACTCTGGAGGCCCTTGACTCTGTGAGAGAGAGCGCAC	1560
QY	1561	TACTACGCTATAGGATAGTAGAGCTGGGGCTGAGGTGTGGAGAAAGAGCGAGGGTCTAC	1620
Db	1561	TACTACGCTATAGGATAGTAGAGCTGGGGCTGAGGTGTGGAGAAAGAGCGAGGGTCTAC	1620
QY	1621	ACCCAAAGTACAAATATTCCTGAATTTGGATCAAGCACACATCAAAAGTGAAGAGGCTTAC	1680
Db	1621	ACCCAAAGTACAAATATTCCTGAATTTGGATCAAGCACACATCAAAAGTGAAGAGGCTTAC	1680
QY	1681	TAA 1683	
Db	1681	TAA 1683	
RESULT 4			
AAC/76693			
ID AAC/76693 standard: cDNA: 3623 bp.			
AAC/76693:			
08-FEB-2001 (first entry)			

Human ORFX ORF2248 polynucleotide sequence SEQ ID NO:4455.

Human: open reading frame; ORX: detection; cytosolic; hepatotropic; villerary; antipsoiatric; antiparkinsonian; neurotropic; neuroprotective; anticonvulsant; osteopathic; antilarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypostensive; dermatological; immunosuppressive; antinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antichryoid; antianemic; gene therapy; cancer; proliferative disorder; hyperextension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antinflammatory disease; coagulation; thrombolysis; contraceptive; SS.

Homo sapiens.

WO200058473-A2.

05-OCT-2000.

31-MAR-2000; 2000WO-US08621.

31-MAR-1999; 99US-0127607.

02-APR-1999; 99US-0127636.

05-APR-1999; 99US-0127728.

30-MAR-2000; 2000US-0540763.

(CURA-) CURAGEN CORP.

Shimkets RA, Leach M:

WPI: 2000-602362/57.

P-PSDB: AAB42484.

Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease -

Claim 5: Page 3679-3681; 5507pp: English.

AAC74446 to AAC74606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; villerary; antipsoiatric; antiparkinsonian; neurotropic; neuroprotective; osteopathic; anticonvulsant; antilarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypostensive; dermatological; immunosuppressive; antinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antichryoid; and antianemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancer, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hyperextension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antinflammatory disease, to enhance coagulation; to inhibit thrombolysis; and as a contraceptive.

Sequence 3623 BP: 848 A; 1005 C; 930 G; 840 T; 0 other:

Query Match 99.7%; Score 1678.2; DB 21; Length 3623;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1680; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 ATGTTCGACGAGATGTGCATTCATCCATTCGCGCTTAATGACCTCGTGGGAAACACA 60

Db 97 ATGTTGCCAGSAGTGTGATCTCCATGTTCTGCTTTAAAGTGTGCTGGTGGGAAAAACACA 156
OY 61 GCGCTGGGGTTCCTCCATGATGTCTTTATTGGAAGCCGTGGACCCAGACTGGACCCCTGAC 120
Db 157 GCGCTGGGGTTCCTCCATGATGTCTTTATTGGAAGCCGTGGACCCAGACTGGACCCCTGAC 216
OY 121 CAGTATGATTACAGCTACGAGAGATTATATCAGAGAAGAACACCACTAGACACTTACC 180
Db 217 CAGTATGATTACAGCTACGAGAGATTATATCAGAGAAGAACACCACTAGACACTTACC 276
OY 181 CAGCTGGGAACCTGAGCTGGTACTACCTAGAGAGACCAAGTGTATCCATGCCAGCCCAAC 240
Db 277 CAGCTGGGAACCTGAGCTGGTACTACCTAGAGAGACCAAGTGTATCCATGCCAGCCCAAC 336
OY 241 CCGCTGGAACAGGTTGGGAGCTGCTGCTCCATGGGAGCACCTTCACATGACAGTGCCTG 300
Db 337 CCGCTGGAACAGGTTGGGAGCTGCTGCTCCATGGGAGCACCTTCACATGACAGTGCCTG 396
OY 301 GCTCCTTTCTCTGGGAATTAAGTGCAGAAAAGTGCAAAATAGTCAAGAGCAACCATGT 360
Db 397 GCTCCTTTCTCTGGGAATTAAGTGCAGAAAAGTGCAAAATAGTCAAGAGCAACCATGT 456
OY 361 GCGCGGGGCCAATGTCTGATTACCCAGAGTCCCTCCCTACTACCGCTGTCTGTAAACAC 420
Db 457 GCGCGGGGCCAATGTCTGATTACCCAGAGTCCCTCCCTACTACCGCTGTCTGTAAACAC 516
OY 421 CCTTACACAGGTCCAGCTGCTCCCAAGTGGTTCCTGTATGACAGGCCCAACCCCTGCCAG 480
Db 517 CCTTACACAGGTCCAGCTGCTCCCAAGTGGTTCCTGTATGACAGGCCCAACCCCTGCCAG 576
OY 481 AATGGGGGACTGCTGCTCCCGGCAATAGCGGAGATCCAGATTCACTGTGCTGCTCCCGAC 540
Db 577 AATGGGGGACTGCTGCTCCCGGCAATAGCGGAGATCCAGATTCACTGTGCTGCTCCCGAC 636
OY 541 CAGTTAAAGGGAATTTCTGGAATAGTTCGATGACTGCTATGTTGGCGATGCTAC 600
Db 637 CAGTTAAAGGGAATTTCTGGAATAGTTCGATGACTGCTATGTTGGCGATGCTAC 696
OY 601 TCTTACCGAGGGAATTAAGTAGAGACAGTCAACAGCATGGTGCCTTCTACTGGAACCTC 660
Db 697 TCTTACCGAGGGAATTAAGTAGAGACAGTCAACAGCATGGTGCCTTCTACTGGAACCTC 756
OY 661 CACCTGCTGTCAGAGAAATTAACAGTCTTATGAGAGAGTGTGAAGCCATGGGATT 720
Db 757 CACCTGCTGTCAGAGAAATTAACAGTCTTATGAGAGAGTGTGAAGCCATGGGATT 816
OY 721 GGGGAACACAAATTTCTGCAAGAACCCAGATGCGAGGAAAAAGCCCTGTATTAATAA 780
Db 817 GGGGAACACAAATTTCTGCAAGAACCCAGATGCGAGGAAAAAGCCCTGTATTAATAA 876
OY 781 GTTACCAATGACAGAGTAATGGTAATGATGATGCTCAGCCTGCTCAGCCAGGAC 840
Db 877 GTTACCAATGACAGAGTAATGGTAATGATGATGCTCAGCCTGCTCAGCCAGGAC 936
OY 841 GTTGTCTACCCAGAGAAAGCCCACTGAGCCATCAACCAAGCTTCGGGGTTTGACTCC 900
Db 937 GTTGTCTACCCAGAGAAAGCCCACTGAGCCATCAACCAAGCTTCGGGGTTTGACTCC 996
OY 901 TGTGGAAGAAGTGAATAGCAGAGAGGAAGATCAGAGATCTGTGAGAGCTTTAAGAC 960
Db 997 TGTGGAAGAAGTGAATAGCAGAGAGGAAGATCAGAGATCTGTGAGAGCTTTAAGAC 1056
OY 961 ACGGGGGGACACCCATGAGAGGCTCCCTCCATGCTCCGTGGGCTTCAGACATCTCC 1020
Db 1057 ACGGGGGGACACCCATGAGAGGCTCCCTCCATGCTCCGTGGGCTTCAGACATCTCC 1116
OY 1021 ATGCCCCAGGGCCACTTGTGTGTGGGGCGCTGATCCACCCCTGCTGAGTCTGCT 1080
Db 1117 ATGCCCCAGGGCCACTTGTGTGTGGGGCGCTGATCCACCCCTGCTGAGTCTGCT 1176
OY 1081 GCGCAGCTGACCGACATTAATAAACGACATCTAAAGCTGTCTAGGGGACCAAGACCTG 1140
Db 1177 GCGCAGCTGACCGACATTAATAAACGACATCTAAAGCTGTCTAGGGGACCAAGACCTG 1236

OY 1141 AAGAAAGAAAGATTTTCATGACGAGACCTTACGGGAGGAAGATATTCAGTACACCCAC 1200
Db 1237 AAGAAAGAAAGATTTTCATGACGAGACCTTACGGGAGGAAGATATTCAGTACACCCAC 1296
OY 1201 TACAAATGAAGAGATGAGATTCGCCCAATGATATTCGATTCAGTTAAAGCCAGTG 1260
Db 1297 TACAAATGAAGAGATGAGATTCGCCCAATGATATTCGATTCAGTTAAAGCCAGTG 1356
OY 1261 GATGTCACACTGCTCTAGAAATCAAAATACGTGAAGACTGTGTTGCTGATGAGTCC 1320
Db 1357 GATGTCACACTGCTCTAGAAATCAAAATACGTGAAGACTGTGTTGCTGATGAGTCC 1416
OY 1321 TTTCCCTCTGGGAGTGAAGTCCACATCTGTGGCTGGGGGTGTACAAACACGAAAGG 1380
Db 1417 TTTCCCTCTGGGAGTGAAGTCCACATCTGTGGCTGGGGGTGTACAAACACGAAAGG 1476
OY 1381 TCCCGCAGCTCTGAGATGCCAAAGTCAAGCTGATTCGCAACACTTGTGCAACTCCGC 1440
Db 1477 TCCCGCAGCTCTGAGATGCCAAAGTCAAGCTGATTCGCAACACTTGTGCAACTCCGC 1536
OY 1441 CAACCTGATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1537 CAACCTGATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596
OY 1501 GGGCAAGACACCTGCAAGGTGACTGTGAGGCCCCCTGACCTGTGAGAAGAGCGCAC 1560
Db 1597 GGGCAAGACACCTGCAAGGTGACTGTGAGGCCCCCTGACCTGTGAGAAGAGCGCAC 1656
OY 1561 TACTAGTCTATGGATAGTACCTGGGCGCTGAGTGTGGGAAGAGGCCAGGGGTCTAC 1620
Db 1657 TACTAGTCTATGGATAGTACCTGGGCGCTGAGTGTGGGAAGAGGCCAGGGGTCTAC 1716
OY 1621 ACCCAGTACCAATTCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1717 ACCCAGTACCAATTCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1776
OY 1681 TAA 1683
Db 1777 TAA 1779

RESULT 5
AAAA4763
ID AAA44763 standard: cDNA; 617 BP.
XX
AC AAA44763;
XX
DT 21-AUG-2000 (first entry)
XX
DE Human secreted expressed sequence tag SEQ ID NO:1338.
XX
KW Human; mouse; chicken; rat; secreted expressed sequence tag; secreted expressed sequence tag; EST; probe; chemotactic; proliferative; immunomodulatory; hematopoietic; chemokine; analgesic; haemostatic; thrombolytic; anti-inflammatory; cytosolic; antibacterial; antifungal; antiviral; antidiabetic; antiparasitic; antiparkinsonian; anticancer; osteopathic; neuroprotective; neurotrophic; antipsoriatic; cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine; autoimmune disorder; multiple sclerosis; allergic condition; insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer; lymphoid cell deficiency; burn; osteoporosis; osteoarthritis; central nervous system disorder; Alzheimer's disease; stroke; Parkinson's disease; Huntington's disease; coagulation disorder; hemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour; infection; depression; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200021991-A1.
XX
PD 20-APR-2000.
XX

PF 15-OCT-1999; 99MO-US24206.
XX
PR 15-OCT-1998; 98US-0104436.
XX
PA (GENE) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR;
XX
DR WPI: 2000-317938/27.
XX
PT Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (SESTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders -
XX
XX
PS Claim 1: Page 549; 803pp: English.
XX
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
CC tissue sources. The SESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemokine; proliferative; immunomodulatory; haematopoietic;
CC chemokine; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC antileukemic; vulnerrary; antitumor; osteoprotective; neuroprotective;
CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antidepressant. The SESTs can be used for gene
CC therapy and in vaccines. The SESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders,
CC Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention.
XX
SQ Sequence 617 BP; 165 A; 169 C; 148 G; 135 T; 0 other;
Query Match 31.3%; Score 526.4; DB 21; Length 617;
Best Local Similarity 99.8%; Pred. No. 2.8e-142;
Matches 527; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 444 GGCCGGGCGCAATGCTTCATTACCCAGAGCTCCCTCACTACCGCTGTGTCTAACAAC 503
QY 421 CCTACACAGGCTCCAGCTCCCTCCAGAGGCTCTGTATGACGGCAACCCCTGGCAG 480
DB 504 CCTACACAGGCTCCAGCTCCCTCCAGAGGCTCTGTATGACGGCAACCCCTGGCAG 563
QY 481 AATGGGCTACCTGCTCCGGCATTAAGCGAGATCCAAAGTTCAACCTGT 528
DB 564 AATGGGCTACCTGCTCCGGCATTAAGCGAGATCCAAAGTTCAACCTGT 611
RESULT 6
AB199281
ID AB199281 standard: cDNA; 397 BP.
XX
AC AB199281:
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:123.
XX
KW Mouse: ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN W020018188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UNIT-) UNIT NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI: 2002-034733/04.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
PS Claim 2: Page 351-352; 2690pp: English.
XX
CC The present invention describes a method for examining ischemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischemic condition (e.g. compressive
CC ischemia, occlusive ischemia or vasospastic ischemia) by measuring
CC expression levels of particular genes (AB199202 to AB199912, encoding
CC the protein sequences in AB57020 to AB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischemic condition-improving
CC drugs or as reagents for ischemic diseases. AB19913 and AB19914
CC represent PCR primers for a mouse ischemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 397 BP; 103 A; 105 C; 111 G; 78 T; 0 other;
Query Match 14.4%; Score 242.6; DB 24; Length 397;
Best Local Similarity 84.7%; Pred. No. 5.2e-60;
Matches 272; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```
|||||
Db 73 CCTTGTGCAACTGCCGACACTATGACCAACCATCATAGTATGATTTGTCG 132
1483 GGAAATCTTGAAAGACCTGGGCAAGACCTGGGAGGTGACCTGGAGCCCTTGAC 1542
133 GGGAACTTCAGAAAGCCGGATCAGACACCTGCGGAGGTGACCTGGAGCCCTTAAC 192
1543 TGTGAGAAAGACGACACTACTAGCTATGAGATAGTGAAGCTGGGCTGAGTGTGG 1602
193 TGTGAGAAAGATGGAATCTACTAGCTACGGGATTTGTAAGCTGGGCTGAGTGTGG 252
Qy 1603 AAGAGCCAGGGGCTTACACCAAGTTACCAATTCCTGATTTGATCAAGCCCATTC 1662
253 AAGAGCCAGGAGCTTACACTAGCTACCAAGTTCTGTAATTTGATAAAGACCAATG 312
Db 1663 AAAAGTGAAGTGGCTTTCAA 1683
313 CACAGCGAGGCTGCTCTGTA 333

RESULT 7
ABAS8323/c
ID ABAS8323 standard; DNA; 451 BP.
XX
AC ABA58323;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #6628.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PR 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX
PT analyzing gene expression in human fetal liver -
XX
PS Claim 1; SEQ ID NO 6628; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
XX
CC measuring human gene expression in a sample derived from human foetal
XX
CC liver. The single exon nucleic acid probes may be used for predicting,
XX
CC measuring and displaying gene expression in samples derived from human
XX
CC foetal liver. The present sequence is a single exon nucleic acid
XX
CC probe of the invention.
XX
CC Note: The sequence data for this patent did not form part of the
XX
CC printed specification, but was obtained in electronic format directly
XX
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
XX
SQ Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

Query Match 7.1%; Score 120.2; DB 22; Length 451;
Best Local Similarity 97.6%; Pred. No. 1.8e-24;
```

```
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 324 TCAGAAAGTGCAAAATAGCTGACAGACCAACCATGTGTGGCGGCAATGTCTCATTC 383
451 TCCTACAGTCCGAAAATAGCTGCAAGACCAACCATGTGTGGCGGCAATGTCTCATTC 392
Qy 384 CCAGAGTCCCTCCCTACTACGCGTGTGTCTGTAACACCCCTTACACAGTCCAGTGTTC 443
391 CCAGAGTCCCTCCCTACTACGCGTGTGTCTGTAACACCCCTTACACAGTCCAGTGTTC 332
Db 444 CCAAG 448
331 CCAAG 327

RESULT 8
AAK06420/c
ID AAK06420 standard; DNA; 451 BP.
XX
AC AAK06420;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 6411.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
XX
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PR 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX
PT brains -
XX
PS Example 4; SEQ ID NO: 6411; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
XX
CC probes which are derived from genomic sequences expressed in the human
XX
CC brain. They can be used to measure gene expression in brain cell samples,
XX
CC which may enable the diagnosis and improved treatment of nervous system
XX
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX
CC epilepsy and cancers. The present sequence is one of the probes of the
XX
CC invention.
XX
SQ Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

Query Match 7.1%; Score 120.2; DB 22; Length 451;
Best Local Similarity 97.6%; Pred. No. 1.8e-24;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 324 TCAGAAAGTGCAAAATAGCTGACAGACCAACCATGTGTGGCGGCAATGTCTCATTC 383
451 TCCTACAGTCCGAAAATAGCTGCAAGACCAACCATGTGTGGCGGCAATGTCTCATTC 392
```

```

QY 384 CCAGAGTCCTCCCTACTACCGCTGTGTCTGTAACACCCCTTACAGAGTCCAGAGTGC 443
Db 391 CCAGAGTCCTCCCTACTACCGCTGTGTCTGTAACACCCCTTACAGAGTCCAGAGTGC 332
QY 444 CCAAG 448
Db 331 CCAAG 327

RESULT 9
AAK32096/C
ID AAK32096 standard; DNA: 451 BP.
AC AAK32096;
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 6653.
XX
KM Human: bone marrow expressed exon: gene expression analysis: probe:
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
PD 09-AUG-2001.
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR:
XX
DR WPI: 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow.
XX
PS Example 4: SEQ ID NO: 6653; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other:

Query Match 7.1%; Score 120.2; DB 22; Length 451;
Best Local Similarity 97.6%; Pred. No. 1.8e-24;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Db 331 CCAAG 327

RESULT 10
AA137947/C
ID AA137947 standard; DNA: 451 BP.
AC AA137947;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #6633 used to measure gene expression in human placenta sample.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
PD 09-AUG-2001.
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR:
XX
DR WPI: 2001-488997/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 6633; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other:

Query Match 7.1%; Score 120.2; DB 22; Length 451;
Best Local Similarity 97.6%; Pred. No. 1.8e-24;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

XX 19-AUG-2002 (first entry)
 DT
 XX
 DE Human genome-derived single exon probe from lung SEQ ID No 6854.
 XX
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurolipomatosis;
 KW tubercous scleriosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haem siderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagazer syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 KW
 XX
 OS Homo sapiens.
 XX
 XX
 PM WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00065.
 PF
 XX
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236539P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SQ, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI: 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples -
 XX
 PS Claim 1; SEQ ID No 6854; 634pp: English.
 XX
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC mRNA, and (b) measuring the labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the labeled nucleic acids bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tubercous scleriosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

CC	haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC	pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic
CC	pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC	and hyaline membrane disease. The present sequence is a single exon
CC	probe of the invention.
CC	Note: The sequence data for this patent did not form part
CC	of the printed specification, but was obtained in electronic
CC	format directly from WIPD at
CC	ftp.wipo.int/pub/published_pcl_sequences.
XX	
SO	Sequence 451 BP: 116 A; 102 C; 127 G; 106 T; 0 other:
QY	Query Match 7.1%; Score 120.2; DB 24; Length 451:
	Best Local Similarity 97.6%; Pred. No. 1.0e-24;
	Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps
Db	
324	TCGAAAGTGTCAAAATACGTGCAGGACACCAATGTGGCCGGGGCCAATGTCTCATTTAC 383
451	TTCTACAGTGCAAAATACGTGCAGGACACCAATGTGGCCGGGGCCAATGTCTCATTTAC 392
OY	384 CCAGAGTCTCTCCCTACTACCGCGTGTCTGTAAACCCCTTACACAGGTCACAGCTGCTC 443
Db	391 CCAGAGTCTCTCCCTACTACCGCGTGTCTGTAAACCCCTTACACAGGTCACAGCTGCTC 352
OY	444 CCAAG 448
Db	331 CCAAG 327
RESULT 12	
AA063951	
ID	AA063951 standard; cDNA; 2033 BP.
XX	
AC	AA063951;
XX	
DF	06-JAN-1995 (first entry)
XX	
DE	Hepatocyte growth factor converting protease coding sequence.
KM	Hepatocyte growth factor; protease; cleavage; active; inactive;
XX	precursor; ds.
OS	Homo sapiens.
XX	
PH	Key Location/Qualifiers
FT	CDS 1..1968
FT	/*tag= a
FT	/product= Hepatocyte growth factor converting
FT	protease.
XX	
PN	EP596524-A.
PD	11-MAY-1994.
XX	
PE	05-NOV-1993; 93EP-0117988.
XX	
PR	05-NOV-1992; 92JP-0296133.
PR	20-NOV-1992; 92JP-0312234.
PR	20-NOV-1992; 92JP-0312242.
XX	
PA	(SHIM.) SHIMOMURA T.
PA	(MITU) MITSUBISHI KASEI CORP.
XX	
PI	Kitamura N, Miyazawa K, Morimoto Y, Shimomura T;
XX	Yamada K;
XX	
DR	WPI: 1994-152921/19.
XX	P-PSDB; AAK3962.
PT	Hepatocyte growth factor converting protease and precursor and
PT	gene encoding them - for producing active two chain HGF from
PT	inactive single chain HGF
XX	

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/lished_pct_sequences.

XX Sequence 2036 BP; 339 A; 743 C; 623 G; 331 T; 0 other;

Query Match 7.0%; Score 117.6; DB 24; Length 2036;
Best Local Similarity 49.8%; Pred. No. 2.2e-23;
Matches 607; Conservative 0; Mismatches 359; Indels 54; Gaps 10;

```

469 AACCCCTGCGAATGGGCTACCTGCTCCCGGCAATAAGCGGAGATCCAGTTACACTGT 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 ACCCTTCCCTGAGAGGGGACACTCCACCTGATCGGGACACGGAGACACCGTGT 804

529 GCGTCTCCGACACGTTCAAGGGGAATTCGTGAAATAGTTCGTGATA---CTGCTAT 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 805 GCGTCTCCGACACGTTGCTGAGCGGCTCTGCAACATCGAGCTGATGAGCGCTGCTTC 864

586 GTTGGCGATGGCTACTCTTACCGAGGGAAATGAAATAGACATGCAACGACATGCGTGC 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 865 TTGGGGAACGGCACTGGGTACTGCGTGGCGGACACCTCAGCTCGGGCTCAGCTGC 924

646 CTTTACTGGAACCTCCACCTCCCTCTGAGGAGAAATTCACATGTTATGAGAGATGCT 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 925 CTGGGCTGGAACTCGATGCTCTACCGAGAGCTGCACTGAGCTCGTGGGCGCGG 984

706 GAAACCCATGGGATTTGGGGAACAAATTTCTGAGAAACCCAGATGGGAGGAAAGCCG 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 985 GCGCTGTGGGCTGGGGCCCGCCATGCTACTGCGGAATCCGGAATCCGAGAGAGGCC 1044

766 TGTGCTTTAATAAGTTACCATGACAGGGAATGGAATTCGTGATGCTCAGGC 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1045 TGTGCTGACGT---GGTGAAGAGAGCGGCTCTCTGAGAGTCTGCTCCCTCGAGGCC 1101

826 TGCTAGCCCAAGAGCTTGGCTACCCAGAGAAAGCCCACTGAGCCATCAACCAAGCTT 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1102 TGGGAATCCCT-----CACCAAGTCCAACTGTACCGGATCTCTGGCGACCTCG 1152

886 CCGGGCTTTGACTCCTGTGGAAGAGCTGAGATAGCAGAGGAGATCAGACATCTAT 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1153 CCTGAGCCAGCTCTCCCGGGGCGGAGCCCTGCGGAGAGGACCAAGAGAGAGCTTC 1212

946 GGAAGCTTTAAGAGACGGCGGCGCAACCCATGCGAGCGGCTCTCAGTCTGCTG 1005
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1213 CTGCGGCCACGATCATGGGCGGCTCTCTCTGCTGCGGCTGCGACACCTG----- 1265

1006 CCTGTGACCATCTCCAGCCCGGCGCACTTCTGTGGGGGCGTGTATCCACCCCTGC 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1266 --GCTGGCGGCATTTACATGSGGAGAGTTCTGCGCGGAGGCTGTGTCCACACTGC 1323

1066 TGGGTGCTCATGCTGCGGCTGACCGACATTAATAAACCAAGCATTTA-----AGG 1119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1324 TGGGTGCTGCTGCGGCGGCGGCTGCTTCTCCACAGCCCGGAGAGAGGCTCTCGG 1383

1120 GTGCTAGGGGACACGACCTGAAGAAGAAATTTATGACAGAGGTTAGGGTGGAG 1179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1384 GTGCTAGGGGACACCTTTTAACCGAGAGAGAGAGTACCGACACCTTGGGATGGAG 1443

1180 AAGATATTCAGATGACGACCTACATGAAGAAAGATAGAGATTTCCCAATGATATGCA 1239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1444 AAGTATATCCCGTACACCTGACTCGGTTCACACCCAGGAGACAC---GACCTCGTC 1500

1240 TTGCTCAAGTTAAAGCAAGTGATGTCATGTCCTGATTCCAATACAGTGAAGACT 1299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1501 CTGATCCGCGTAAAGAAAGAGGAGGACCGCTGTGCGACACGCTCGCAATTTGTGACGCC 1560

1300 GTGTGCTTGCGCTGAT-----GGGTCTTTTCCCTGTGGGAGTGAAGTGCACATCTGCG 1353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1561 ATCTGCTCGCCGAGCCGCGAGACCTTCCCGCGAGGACACAAATGTCGCGGCG 1620

1354 TGGGG-----TGTTACAGAAACAGGAAAGGGTCCCGGACACTCTCTGATGCGCAAGTGC 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1621 TGGGGCACTTGATGAGAACGTGACGGGCTACTCCAGCTCCCTCGGGAGGCCCTGTGTC 1680
```

```

Oy 1408 AACCTATTTGCCAACACTTGTGCACTCCCGCAACTGTATGACACATGTTGATGAC 1467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1681 CCCCTGTGCGGACACCAAGTGGAGGCTGTGAGGCTTAAAGGCGCGGACATCAGCCCC 1740

1468 ACTATGATCTGTGACAGAAATCTTCAGAAACCTGGGCAAGACACCTGCCAGGGTGTACT 1527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1741 AACTATCTGTGTGCGGCTACTT---CGACTGCATGTCGACGCTGCGAGGGGAGCTCA 1797

1528 GAGAGCCCCCTCACTGTGAGAAAGAGGACCTACTACTGCTATAGGATAGAGAGCTGG 1587
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1798 GGGGGGCCCCCTGTGCGGAGAGAGAGCGGTGCTTACTCTTACGCATATCAGCTGG 1857

1588 GCGCTGAGTGTGG-----AGAGGCGAGGGTCTACACCCAGTAATCCAAATTCCTG 1641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1858 GTGAGCGGCTGGGGGCGGCTCCACAAAGCCGGGGGTCTACACCCGCTGGCAACTATGTG 1917

Oy 1642 AATTGATCAAGCCACCAT 1661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1918 GACTGATCAACGACCGGAT 1937

RESULT 14
ABL65242
ID ABL65242 standard; DNA; 2036 BP.
XX
AC ABL65242;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3579.
XX
KW Human; cancer; colon; breast; ovary; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Kilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN W0200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
```


Db	1045	TGGTGTCTACGCT----	-GGTGAAGAGACAGCCGGCTCTCTCTGGAGTACTGCCCGCTGTGAAGGCC	1101
Qy	826	TGCTTCAGCCAGGACGCTTGCC	TACCCAGAGAAAGCCCACTGAGCCATCTCAACCCAGCTT	885
Db	1102	TGCGAATTCCT-----	-CACAGAGTCCAACTGTACACCGGATCTCTGTGGAGACCTTG	1152
Qy	886	CCGGGGTTTTACCTCTGTGGAAAGACGTAGACAGAGAGCAATCAAGACATCTAT		945
Db	1153	CCCGAGGACACTCTCCCGGGGCGCAGGGCTCTGGGACAGAGAGGACAAAGAGACGCTTC		1212
Qy	946	GGAGGCTTTAAGACACAGCGCGGCGAAGCACACCATGACAGCGCTCCCTCACTCTCTGCTG		1005
Db	1213	CTGGCGGACAGCTTTCATATGGGCGGTCTCTCTGCTGCGCGGCTGCACCCCTG-----		1265
Qy	1006	CCCTTGACATCTTCATGCGCCCAAGGGCGCACTTGTGTGGGGCGCTGATCCAGCCCTGCG		1065
Db	1266	--GCTGGCGGCGCATCTACATGGGGAGACACTCTTGCGCGGGAGCGCTGTTCACACACTTCG		1323
Qy	1066	TGGGGTCACTCTCCCGCCACTGACCGCATATATAAAACGAGACATCTTA-----	AGG	1119
Db	1324	TGGGTGGTGTCTGGCGCGCCACACTGTCTTCCCAAGCCCCCGGAGAGACAGCTCTCCGCG		1383
Qy	1120	GTGCTAGGGGACAGGACCTTCAAGAAAGAAATTTACATGACAGAGCTTTAGGCTGAG		1179
Db	1384	GTGCTGGGCGACGACACTTCTTAAACCCGACGACGAGCGATGACGAGACACTTGTGGATCGAG		1443
Qy	1180	AAGATATTCTCAAGTCAAGCCAGCCACATGGAAGAGATGAGATTTCCCAACATGATATTGCA		1239
Db	1444	AAGTACATCCCGACACACCTGTACTCGGTGTTCAAACCCAGGACCAAC--GACCTGCTGC		1500
Qy	1240	TTCGCAAGTTTAAAGCAGTGGATGGTCACTGTGCTGCTAGCAATCCCAATATGGAAGACT		1299
Db	1501	CTGATCCGCGCTGAAGAAAGGAGGAGACCGCTGTGCACACGCTCGCACTTGTGTGACGCC		1560
Qy	1300	GTGTGTCTTGCCCTGAT-----	GGGTCCCTTTCCTCTGTGGAGTGAGTCCACACTCTTGCG	1353
Db	1561	ATCTGCTCCGCTCCGAGCGCGGAGBACCTTTCCTCCGAGGACACAAAGTGGCAATTTGGCGGC		1620
Qy	1354	TGGGG-----	TGTTACAGAAACAGSAAAGGGGTCCCGGACGCTCTGGATGTCCAAAGTC	1407
Db	1621	TGGGGACCACTTGGATGAGAAAGTACGAGCGGCTACTTCAGTCCCTGTGGGAGGCGCGTGTGC		1680
Qy	1408	AAGCGATTCGCAACACTTTGTGTGAATCTCCCGCACACGTATGACCACTATGATTTGATGAC		1467
Db	1681	CCCCCTGGTCCGACACACAAGGTGAGACCTTGAGGCTCTAGCGGGCGCGACATACGCCCT		1740
Qy	1468	AGTATGATCTGTCCAGCAAAATCTTTCAGAAACCTTGCGGAGACACCTGCGACGGTGACTCT		1527
Db	1741	AACATGCTCTGTCCCGCTTACTT-----	CGACTGCAAGTCCACAGCCCTGCGCAAGGGGACTGA	1797
Qy	1528	GGAGGCCCTCCGACGTGGAGAAAGAGCGCACACTACTAGTCTATGAGGATAGTGAACCTGG		1587
Db	1798	GGGGGGCCCTCTGGCGTGGAGAAAGAAAGGCGTGGGTACTCTACGGGATATCTACACTGG		1857
Qy	1588	GGCTCTGAGAGTGTGG-----	AAGAGCCAGGGGGTCTACACCCAGTATCCAAATTTCTCG	1641
Db	1858	GGTGAAGGCTCTCGGGGGGCTCCACAAGACCGGGGGTCTACACCGCGGTGGCAACTATGTGG		1917
Qy	1642	AATTTGATCAAAAGCCACCAT	1661	
Db	1918	GACTGGATCAACAGCCGAT	1937	
RESULT 15				
ABAT0919/C				
ID ABAT0919 standard. DNA: 117 Bp.				
XX ABAT0919:				
XX AC				
XX 01-FEB-2002 (first entry)				
XX				
XX Human foetal liver single exon nucleic acid probe #19224.				
XX				

XX	Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
KX	Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX	Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
OS	Homo sapiens.
XX	MO200157277-A2.
PN	09-AUG-2001.
PD	30-JAN-2001; 2001WO-0500669.
XX	04-FEB-2000; 2000US-0180312.
XX	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632666.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236559.
PR	04-OCT-2000; 2000GB-0024263.
XX	(MOLE-) MOLECULAR DYNAMICS INC.
PA	Penn SG, Hanzel DK, Chen W, Rank DR;
PI	WPI; 2001-483447/52.
DR	Human genome-derived single exon nucleic acid probes useful for
XX	analyzing gene expression in human foetal liver.
XX	Claim 4: SEQ ID NO 19224; 639pp + sequence listing; English.
XX	The invention relates to a single exon nucleic acid probe for
CC	measuring human gene expression in a sample derived from human foetal
CC	liver. The single exon nucleic acid probes may be used for predicting,
CC	measuring and displaying gene expression in samples derived from human
CC	foetal liver. The present sequence is a single exon nucleic acid
CC	probe of the invention.
CC	Note: The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct-sequences.
XX	Sequence 117 BP; 24 A; 24 C; 40 G; 29 T; 0 other;
XX	Query Match 7.0%; Score 117; DB 22; Length 117;
XX	Best Local Similarity 100.0%; Pred. No. 8e-24;
XX	Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
Oy	332 TGCAGAAATCGCGACAGACACACCCATGCGCGGGGCGCATGTCATTACCCAGAGTC 391
Db	117 TGCAGAAATCGCGACAGACACACCCATGCGCGGGGCGCATGTCATTACCCAGAGTC 58
Oy	392 CTCCTACTACCGCTGTGTCTGTAAACACCCCTTACACAGSTGCCAGCTGCTCCAG 448
Db	57 CTCCTACTACCGCTGTGTCTGTAAACACCCCTTACACAGSTGCCAGCTGCTCCAG 1

Search completed: March 6, 2003, 16:01:34
Job time : 353.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 15:20:13 : Search time 2178.5 Seconds
(without alignments)
12511.822 Million cell updates/sec

Title: US-09-912-559-1

Perfect score: 1683
Sequence: 1 atgttgcagatgtctga.....aaagtgaagtgcttctaa 1683

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estopl:*
6: em_estro:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est12:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	698.2	41.5	891	13	BI761782 603046775
2	451.6	28.8	592	10	BE032018 BE032018
3	447.2	26.6	802	13	BC972579 602841165
4	437.8	26.0	674	9	AA217892 mv54b08.r
5	435.4	25.9	751	13	BI148082 602912432
6	435.2	25.9	570	13	BM508620 1137h03.y

7	416.4	24.7	929	12	BF788188	602113411
8	396.4	23.6	791	13	BI332440	602880841
9	395.8	23.5	546	9	AA237499	mA10B10.r
10	392.6	23.3	608	10	AV601564	AV601564
11	390.4	23.2	902	12	BF780971	602105493
12	386	22.9	682	13	BI220028	602934913
13	384.6	22.9	929	12	BF785781	602112402
14	372.4	22.1	969	12	BF384535	602046804
15	370.6	22.0	654	10	AMA75402	60257502
16	370.4	22.0	587	13	BC972681	602839029
17	365.6	21.7	659	10	AM610902	602114406
18	363.6	21.6	594	13	BM503097	602103735
19	348.8	20.7	716	12	BF789705	602103735
20	335.8	20.0	655	10	BB569555	BB569555
21	330.8	19.7	934	13	BI765113	603051314
22	327.8	19.5	345	14	T68666	yC43e09.r
23	319	19.0	486	9	AA268125	vB08612.r
24	311.8	18.5	600	12	BF788541	602114406
25	304	18.1	335	14	R89458	yQ01a10.r
26	293.2	17.4	304	14	T77362	yD72g02.r
27	279.8	16.6	521	13	BI221626	602836942
28	275.8	16.4	912	13	BC972626	602836922
29	275.6	16.4	366	14	T93666	yE06a11.r
30	266	15.8	434	14	T84369	yD37e04.r
31	258.6	15.4	781	12	BF533788	602075308
32	254.6	15.1	458	14	R10295	yF36e12.r
33	245.4	14.6	719	13	BI102281	60288594
34	242.8	14.4	593	9	A1035669	uD60b09.x
35	242.6	14.4	384	14	M66981	me29h02.r
36	242.6	14.4	397	9	AA002504	mg45a03.r
37	241.8	14.4	513	13	BI247736	602959379
38	241.4	14.3	353	14	R11014	yF38g10.r
39	235.6	14.0	489	9	A1787532	u119e04.y
40	235.4	14.0	393	14	T50880	yB88g07.r
41	215	12.8	530	10	BE235171	142539 MA
42	204.8	12.2	361	9	AA108230	mp38a11.r
43	199.4	11.8	320	14	T69534	yD12f11.r
44	154.6	9.2	642	13	BJ067550	BJ067550
45	150.4	8.9	810	9	A1255992	u176c01.x

ALIGNMENTS

RESULT 1
BI761782 891 bp mRNA linear EST 25-SEP-2001
DEFINITION 603046775F1 NIH_MGC_116 Homo sapiens CDNA clone IMAGE:5187066 5',
LOCUS mRNA sequence.
ACCESSION BI761782
VERSION BI761782.1 GI:15753360
KEYWORDS EST.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clome distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM11467 row e column: 19
High quality sequence stop: 839.
Location/Qualifiers 1..891

FEATURES

source
1..891

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5187066"
/clone_1id="NIH_MGC_116"
/lab_host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC library."

```

BASE COUNT 243 a 234 c 211 g 203 t

Query Match 41.5%; Score 698.2; DB 13; Length 891;
 Best Local Similarity 97.2%; Pred. No. 1.4e-187;
 Matches 775; Conservative 0; Mismatches 13; Indels 9; Gaps 6;

```

Db 1 ATGTTGGCCAGATGCTGCATCTCATGTTCTGTTAATGGCTGTGGGAAAGACA 60
81 ATGTTGGCCAGATGCTGCATCTCATGTTCTGTTAATGGCTGTGGGAAAGACA 140
61 GCCTGGGTTCTCCCTGATGTCCTTTATTGGAAGCTGGACCCAGACTGGACCCCTGAC 120
141 GCCTGGGTTCTCCCTGATGTCCTTTATTGGAAGCTGGACCCAGACTGGACCCCTGAC 200
121 CAGTATGATTACAGTACGAGATTTAATCAGAGACACACAGTACAGACACTTACC 180
201 CAGTATGATTACAGTACGAGATTTAATCAGAGACACACAGTACAGACACTTACC 260
181 CAGTATGATTACAGTACGAGATTTAATCAGAGACACACAGTACAGACACTTACC 240
261 CAGTATGATTACAGTACGAGATTTAATCAGAGACACACAGTACAGACACTTACC 320
241 CCCTGGACACAGTGGGGAGTGGCTGGTCCATGGGAGACCTTACATGAGAGTGGCTG 300
321 CCCTGGACACAGTGGGGAGTGGCTGGTCCATGGGAGACCTTACATGAGAGTGGCTG 380
301 GCTCTCTTCTCGGAAATAGTGTACAAAATGCAAAATACGTGCAAGAGACAAACCATGT 360
381 GCTCTCTTCTCGGAAATAGTGTACAAAATGCAAAATACGTGCAAGAGACAAACCATGT 440
361 GCGCGGGGCAATGTCATATACCAAGAGTCCCTTACTACCGCTGTGTGTTAAACAC 420
441 GCGCGGGGCAATGTCATATACCAAGAGTCCCTTACTACCGCTGTGTGTTAAACAC 500
421 CATTACACAGTCCAGCTCTCCCAAGTGGTCTGTATGAGAGCAAAACCCCTGCCAG 480
501 CATTACACAGTCCAGCTCTCCCAAGTGGTCTGTATGAGAGCAAAACCCCTGCCAG 560
481 AATGGGCTACTGCTCCCGCATTAACGGAGATCCCAATTAACCTGGCGCTGCCGAC 540
561 AATGGGCTACTGCTCCCGCATTAACGGAGATCCCAATTAACCTGGCGCTGCCGAC 620
541 CAGTTAAGGGGAAATTTCTGAAATAGTTCATGATGACTGCTATGTTGGAGTGCATC 600
621 CAGTTAAGGGGAAATTTCTGAAATAGTTCATGATGACTGCTATGTTGGAGTGCATC 680
601 TCTTACCGAGGAAATATAGAGACAGTC -AACCGACATGGTGGCTTTACAGGAATC 659
681 TCTTACCGAGGAAATATAGAGACAGTC -AACCGACATGGTGGCTTTACAGGAATC 740
660 CC-ACCTCTCTTGCAGAGAAATTAACAATGTTATGAGAG -ATGCTGAAGAACCATG 715
741 CCAAGCTCTCTTGCAGAGAAATTAACAATGTTATGAGAGAGATGCTGAAGAACCATG 800
716 GGAATTTGGGAA--CACATTTCTGCAG-AAACCCAGATGGCAGG-AAAACCCCTGTGTC 771
|||||

```

```

Db 801 GGAATTTGGGAAACACCAATTTCTGCAGAGAAACCCAGATGGCCAGGAAAAAGCCCTGTGC 860
QY 772 TTTATTAAGTTACCAA 788
Db 861 TTTATTAAGTTACCAA 877

```

RESULT 2
 BE032018 592 bp mRNA linear EST 09-JUL-2000
 LOCUS
 DEFINITION
 ACCESSION BE032018
 VERSION BE032018.1 GI:8327027
 KEYWORDS
 SOURCE
 ORGANISM
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt. trimmed with phred
 v0.980904.e. Vector identified by cross-match with the -minscore 18
 and -mismatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCATGATGACCAT
 BACKWARD: GTTTCACAGTACAGAG
 Plate: 69 row: K column: 12
 Seq primer: ATTGAGTACATATG
 location/qualifiers

FEATURES
 source

1..592
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_1id="MARC_1PIG"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
 library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."

BASE COUNT 144 a 152 c 174 g 122 t

Query Match 26.8%; Score 451.6; DB 10; Length 592;
 Best Local Similarity 86.0%; Pred. No. 2.2e-117;
 Matches 518; Conservative 0; Mismatches 69; Indels 15; Gaps 1;

```

QY 522 CAGCTGAGCTGTCCTCCGACCACTTCAAGGGAATTTCTGGAATAGTTCTGATGACTG 581
Db 6 CAGCTGAGCTGTCCTCCGACCACTTCAAGGGAATTTCTGGAATAGTTCTGATGACTG 65
582 CTATGTTGGGAGATGCTACTCTTACCGAGGGAATTAATAGGACAGTCAACAGCATGC 641
6 CTATGTTGGGAGATGCTACTCTTACCGAGGGAATTAATAGGACAGTCAACAGCATGC 125
642 GTGCTTACTGSAATCCCACTCTCTTGCAGAGAAATTAACAGATGTTATGAGAGA 701
Db 126 GTGCTTACTGSAATCCCACTCTCTTGCAGAGAAATTAACAGATGTTATGAGAGA 185
702 TGCTGAACCAATGAGGATTTGGGAGAAATTTCTGAGAGAAACCCAGATGGAGAGAAA 761
Db 186 TGCTGAACCAATGAGGATTTGGGAGAAATTTCTGAGAGAAACCCAGATGGAGAGAAA 245
762 GCCCTGTGCTTTATTAAAGTTACCAATGACAGGTGAATGGGAATGATGATGTCCTC 821

```

D6		BG972579 LOCUS	802 bp mRNA linear EST J2-JUN-2001		
246	GCCCGTGGTTTCATTAAAGTGACCAATGTGAAGCTACTGCACCCTC			305	
OY	AGCCATGCCACCCAGACGTTGCTACCCAGAGAAGCCCCACTGAGCATCACCAA			881	
D6	TGCTGCTGCAC-----CCAGAGGAGAGCCCATTAGAAGCTTGACCCA			350	
OY	GGTTCGGGGTTTTACTCTGTGGAAAAGACTGAGATAGCAGAGAGAGATCAAGACAAT			941	
D6	GCCTCCCGGGTTTAGAAGTGTGGAGGACAGACAGCAAGAAAGAGTTCAAGAGAT			410	
OY	CATTAAGAGCTTTAAGAGCAGCGGGGCGAACACCATAAGCAGCGTCCCTCCAGTCTCC			1001	
D6	CTPACGAGGCTTTAAGAGCAGCGGGGCAAGACCCCTGGCAGGCGTCCCTCACAGCTTC			470	
OY	GCCTGCTCTGACACTCCATCCATGCCCCAGAGGCGCACTCTGTGGTGGGGCGCTGATCCACC			1061	
D6	CTTGCATCATCGACCTTCATGCCCCAGAGGCGCACTCTGTGGGGGCGCGTATCCACC			530	
OY	CTGCGCGGTGCTCACCTCCCTGCCACTGCACCGACATTAAGACAGACATCTAAAGGTGT			1121	
D6	CTGCTGGGGTGTCAGCCCTCCATTGCACGACATTAAGCCAAATATCTMAAAGATGT			590	
OY	1122 GC 1123				
D6	591 GC 592				
RESULT 3					
BG972579					
DEFINITION	602841165F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4975384				
ACCESSION	BG972579				
VERSION	BG972579.1 GI:14360216				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL COMMENT	Unpublished (1999) Contact: Robert Stransberg, Ph.D. Email: cgabsf@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.jnl.gov/ Plate: LLMH0967 row: a column: 17 High quality sequence stop: /85. Location/Qualifiers				
FEATURES					
source	1..802	/organism=“Mus musculus” /strain=“FVB/N” /db_xref=taxon:10090” /clone=“IMAGE:4975384” /clone_1lb=“NCI_CGAP_Kid14” /_lab_host=“DHIOB (T1 phage-resistant)” /_note=Organ: kidney; Vector: pCMV-Sport5; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: This is a NCI-CGAP library.- ”			
BASE COUNT	211 a 230 c 203 g 158 t				
ORIGIN					
Query Match	26.6%; Score 447.2; DB 13; Length 802;				
Best Local Similarity	77.7%; Pred. No. 4.6e-116;				
Matches 580; Conservative	0; MisMatches 158; Indels 8; Gaps 3				

[illegible]


```

Db 363 AGTATGATTTTGGGGGAACTTCAGAAACCCCGGATCAGACACCTTCCAGGATGCTAC 422
QY 1528 GGAGGCCCTGACCTGTGGAAGAGAGCGACCTACTAGCTTATGGAGTATGAGCTGG 1587
Db 423 GGGGGCCCTCAACCTGTGAGAGGATGGAACTTACTACCTACGCGGATTTGAGCTGG 482
QY 1588 GGCCTGAGTGTGGAGAGAGCGAGGAGTCTACACCCAGTTACCAATTCCTGATTTGG 1647
Db 483 GGCAGAGATGTGGAGAGAGCGAGGAGTCTACACTCAAGTACCAAGTTCCTGATTTGG 542
QY 1648 ATCAAGCCACCATCAAAAGTGAAGTGTG 1675
Db 543 ATAAAGCCACCATGACAGAGAGGCTGG 570

RESULT 7
BF78188 929 bp mRNA linear EST 12-JAN-2001
LOCUS 60211341P1 NCI_CGAP_K1d14 Mus musculus cDNA clone IMAGE:4241642
DEFINITION 5', mRNA sequence.
ACCESSION BF78188
VERSION BF78188.1 GI:12093224
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS 1 (bases 1 to 929)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M9858 row: e column: 03
High quality sequence stop: 686.
Location/Qualifiers
1..929
location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1db="IMAGE:4241642"
/lab_host="NCI_CGAP_K1d14"
/note="Organ: Kidney; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI/CGAP Library."
BASE COUNT 238 a 275 c 253 g 162 t 1 others
ORIGIN
Query Match 24.7%; Score 416.4; DB 12: Length 929;
Best Local Similarity 75.9%; Pred. No. 3e-107;
Matches 529; Conservative 0; Mismatches 162; Indels 6; Gaps 1;

QY 1 ATGTTGCCAGAGATGCTGATCTCCATGTCCTGTAATGCGTGTGGGGAAGACA 60
Db 44 ATATTTGTCAGATGATGTTCCGTCCTGCTGCTAATGCGCTGGGGGAAGTGA 103
QY 61 GCCTGGGCTTCCCTCATGTCCTTATTGGAAGAGCTGGACCCAGACTGACCCCTGAC 120
Db 104 GTATGGGCTCTCAGCTGATGTCCTTCATGTCGCCGCCAGACCATNGACCCCGCAT 163
QY 121 CAGTAAATTAACAGTACAGAGATTATATCAGAGAGAACACCAAGAGACACACTTACC 180
Db 164 GACTATTACTACAGCTATGAGACGTCAGCCAGACAGAACCCCAAGTGTACAGCAGACC 223

```

```

QY 181 CATGTGAGAAATTCCTGATGTAACATACATGAGACCAACGATCATCCAGGCCAAC 240
Db 224 ACCCTGAGAAACCCGAGCTGTACT-----ATGAAAGCATATATCATCCAGCTCCAC 277
QY 241 CCCTGTGAACAGGCTGGGAGCTGCTCTGTCATGSGAGCACTTTCACATGATGCTGCTG 300
Db 278 CCCTGTGAACAGGCTGGGAGCTGCTCTGTCATGAGAGGAGATCTCTTCAGTTGCAAGTCCCA 337
QY 301 GCTCTTCTCTGAGAAATTAAGTGTGCAAAAGTACGTCGACGACCAACCACTGT 360
Db 338 GCGCCCTTCTGAGGAGCGGTCGCCAGACTGACAAAACAAAGTGAAGACCAACCACTGT 397
QY 361 GCGCGGCGCAATGTCTCATTTACCCAGAGCTCTCTCTACATCCGCTGTCTGTAAACAC 420
Db 398 GTTCATGATGATGCTCTCATTTACCCAGAGCAACCCCTACTACGCTGTGCTCAAAATAC 457
QY 421 CCTTACAGAGGTCCAGCTGCTCCCAAGTGTGCTCTGTATGACAGGCAACCCCTGCGAG 480
Db 458 CCTTACAGGAGACAGACTGCTCCAAAGTGTCTCGGATGACAGGCCAAACCCCTGCGAG 517
QY 481 AATGGGCGTACCTGCTCGCGCATTAAGAGATGCAAGTTCACTGCTGTGCTGCGAG 540
Db 518 AATGGGCGATGCTGTCTCCAGACAGACGAGATCCAGTTTACTGCTGCTGCTGACAG 577
QY 541 CAGTTCAAGGGAATTTCTGTGAATTAAGTGTGATGACTGTCTATGCTGCGATGCTAC 600
Db 578 CAGTAAAGGGGAATTTCTGAAATAGGTCGAGCTGTATGTGCTGATGATGCTAC 637
QY 601 TCTTACCGAGGAAATGAATGAGACAGTCAACACAGCTGCGCTTACGAGAACCC 660
Db 638 TCTTACCGAGGAAATGAATGAGACAGTCAACACAGCTGCGCTTACGAGAACCC 697
QY 661 CAGCTCTCTTGTGAGAGAAATTAACATGTTTATG 697
Db 698 CAGCTCTCTGAGGAGACTATACATGATGCTGCG 734

RESULT 8
BF132440 791 bp mRNA linear EST 30-JUL-2001
LOCUS 60298084P1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5133587 5',
DEFINITION mRNA sequence.
ACCESSION BF132440
VERSION BF132440.1 GI:15017097
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS 1 (bases 1 to 791)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M1328 row: a column: 12
High quality sequence stop: 787.
Location/Qualifiers
1..791
location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1db="IMAGE:5133587"
/lab_host="NCI_CGAP_L19"
/note="Organ: liver; Vector: PCMV-SPORT6; Site 1: NotI;

```


Site 2: Salt: Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: this is a NCI-CGAP library."

BASE COUNT 202 a 234 c 194 g 161 t
 ORIGIN

Query Match 23.6%; Score 396.4; DB 13; Length 791;
 Best Local Similarity 75.6%; Pred. No. 1.4e-101;
 Matches 520; Conservative 0; Mismatches 161; Indels 7; Gaps 2;

OY 1 ATGTTGCGGATGCTGATCTCATGTCGCTGTTATGCTGCGGAAACCA 60
 || ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 23 AATATTGACAGATGTGGTGTCCGTGCTGCTGCTATGCTGCGGAAACCA 82
 OY 61 GCCTGGGTTCTCCCTGATGCTTTATTGAAAGCTGACCCAGACCTGAC 120
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 83 GTCATTGGGCTTCACAGATGCTTCATGCGCCCGGAGAGCCAGATTGACCCCGAT 142
 OY 121 CAGTATGATTACAGTACGAGATTTATATGAGAGAGACACAGTACACTTACC 180
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 143 GACTATTACTACAGCTATGAGAGTCCAGCCAGAGACAGACCCAGTGTACAGGAGAC 202
 OY 181 CATGTTGAGATCTGCTACTACACTGAGAGACCAAGCTGATCCAGCCGAC 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 203 ACCCTGAGAACCCCGACCTGTGCTACT-----ATGAGAGCATGATCTCATGCTGACAC 256
 OY 241 CCCTGTGACACGCTGGGAGCTGCTCCATGAGGACACCTTTCATGCTGCTG 300
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 257 CCCTGTGACACGCTGGGAGCTGCTCATGAGAGGATACCTTCACTGATGCTGCTG 316
 OY 301 GCTCTTTCTCTGGGATTAAGTGTGCAAAATGCAAAATGCTGCAAGGACAAACCATCT 360
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 317 GCCCCTCTCTGGGAGCCGCTGCTGACACTGACAAACAAAGTCAAGGACAAACCATGT 376
 OY 361 GGGCGGGGCAATGCTCATTAACCAAGTCTCTCTACCGCTGCTGCTGTAACAC 420
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 377 GTCATGCTGATGCTCTCATTAACCAAGGACACCTCTACTACCGCTGCTGCTGCAAAATAC 436
 OY 421 CCTTACACAGTCTCCAGCTCTCCCAAGTGTCTCTGATATGACGCGCAAAACCCCTGCGAC 480
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 437 CCTTACACAGGACGACGCTCTCCCAAGTGTCTCCGATGACGCGCAAAACCCCTGCGAC 496
 OY 481 AATGAGGCTTACCTGCTCCGCAATAGCGAGATCCAAATTCACCTGCTGCTGCGAC 540
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 497 ATGGCGGAGCTGTCTCCGCAACAGCGAGATCCAGGTTTACCTGCTGCTGCGAC 556
 OY 541 CAGTCAAGGAGAAATGTAATAGACAGTCAACCAAGCATGCTGCTTACTGAGACTC 599
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 557 CAGTCAAGGAGAAATGTAATAGACAGTCAACCAAGCATGCTGCTTACTGAGACTC 616
 OY 600 CTCTTACGAGGAGAAATGTAATAGACAGTCAACCAAGCATGCTGCTTACTGAGACTC 659
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 617 CTCTTACGAGGAGAAATGTAATAGACAGTCAACCAAGCATGCTGCTTACTGAGACTC 676
 OY 660 CCAGCTCTCTTGCAGGAGAAATTAAC 687
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 677 CCAGCTCTCTTGCAGGAGAAATTAAC 704

RESULT 9
 AA237499 546 bp mRNA linear EST 03-MAR-1997
 LOCUS mx10b10.1 Soares mouse NML Mus musculus cDNA clone IMAGE:679771 5'
 DEFINITION similar to TR:G1345398 G1345398 HGF ACTIVATOR LIKE PROTEIN.; mRNA
 sequence.
 ACCESSION AA237499
 VERSION AA237499.1 GI:1861520
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 546)

AUTHORS

Matta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Knobbs, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE

The WashU-HMNI Mouse EST Project
 Unpublished (1996)
 Contact: Maria M/Mouse EST Project
 WashU-HMNI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

COMMENT

Email: mouseest@wustl.edu
 This clone is available royalty-free through LBNL: contact the
 IMAGE Consortium (InfoImage.lbnl.gov) for further information.
 MGI:419475

FEATURES

Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 405.
 Location/Qualifiers
 1..546
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:679771"
 /clone_lib="Soares mouse NML"
 /tissue_type="Liver"
 /lab_host="DH10B"
 /note="vector: pT73D-Pac (pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTCACCATCTGTAAGTGGAGCGCGCCGATCTGTTTCTTTTCTTTT 3'];
 double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector library
 constructed and normalized by Bento Soares and M.Felina
 Bonaldo."

source

BASE COUNT 141 a 142 c 151 g 112 t
 ORIGIN

Query Match 23.5%; Score 395.8; DB 9; Length 546;
 Best Local Similarity 83.1%; Pred. No. 1.7e-101;
 Matches 451; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

OY 868 GAGCATCAACCAAGTCTCCGGGTTGACTCGTGGAAAGAGTACAGAGAGG 927
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 4 GAGCTGTGATGAGAGCTGCCAGAGCTTCAGTCTGAGAGAAAGAGAGTACAGAAAC 63
 OY 928 AAGATCAAGAGATCTATGAGAGCTTTAAGACAGCGCGGCAACACCCATGCGAGCG 987
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 64 GCAGTCAAGCGTATCTAAGGAGGCTTTAAGACAGACAGCAACACCCGTCGAGGTG 123
 OY 988 TCCCTCAGTCTGCTGCTCTGACATTCATGCCCCAGGGCCACTTCTGTGTGGG 1047
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 124 TCCCTCAGACCTCAGCTGCTGCTGACACCTCCATGCCCCAAGGCCACTTCTGTGTGGG 183
 OY 1048 GCGGTATCCACCCCGTGGGGTGTCTACCTGCTCCAGTCCAGGACGAGATTAACAGCA 1107
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 184 GCGGTATCCACCCCGTGGGGTGTCTACCTGCTCCAGTCCAGGACGAGATTAACAGCA 243
 OY 1108 CATCTAAAGTGTGCTAGGAGACCAAGACTGAGAAAGAAATTTTCATGAGCAGAGC 1167
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 244 CATCTAAAGTGTGCTAGGAGACCAAGACTGAGAAAGAAATTTTCATGAGCAGAGC 303
 OY 1168 TTTAGGTGAGAGATATTCAAGTACAGCAATGAAAGAGATGAGATTCCCCAC 1227
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 304 TTTAGGTGAGAGATATTCAAGTACAGCAATGAAAGAGATGAGATTCCCCAC 363
 OY 1228 AATGATATTGATTCCTCAAGTTTAAAGCAGTGGATGCTGCTGCTGATTCCTCAAA 1287
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 364 AATGATATTGATTCCTCAAGTTTAAAGCAGTGGATGCTGCTGCTGATTCCTCAAA 1423
 OY 1288 TAGGTGAAGACTGTGTGCTTCTTCCTCTGAGGAGTGAAGTCCACATC 1347

Accession	Sequence	Position
Db	TATGTAAAGACATGATATTTTGCCACGACCCCTTTCCCTCTGAACTCATGAGGCCATC	483
Db	424	
OY	TCTGAGTGGGATGTTACAGAAACAGGAAAAGGATGCCGCCAGCTCTCTGATGCCAAATC	1407
OY	1348	
Db	TCCTGGTGGGGGATACAGAAACCAAGGAGGATTCGCCAGCTCTCTGATGCCAAATC	543
OY	484	
OY	1408 AAC 1410	
Db	544 AAG 546	

RESULT	10
AV601564	
LOCUS	608 bp mRNA linear EST-27-NOV-2001
DEFINITION	AV601564 Bos taurus kidney fetus Bos taurus cDNA clone EIKI005A04 5', mRNA sequence.
ACCESSION	AV601564
VERSION	AV601564.1 GI:9723884
KEYWORDS	EST.
SOURCE	COW.
ORGANISM	Bos taurus
<i>Eukaryota; Mammalia; Artiodactyla; Bovidae; Ruminantia;</i>	

REFERENCE
AUTHORS
1 (bases 1 to 608)
Takasuga, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H.

TITLE	JOURNAL
Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs	Nucleic Acids Res. 29 (22), E108 (2001)
3120564	

FEATURES

COMMENT

Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odawara, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp

Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers

```

/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_1fb="Bos taurus kidney fetus"
/clone_1lb="Bos taurus kidney fetus"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/notes="Vector: pZRL1, Site-1: SalI; Site-2: NotI; Poly A
was deleted from a NotI site"
BASE COUNT      148 a      171 c      151 g      135 t      3 others
ORIGIN

```

Query Match	23.3%	Score 392.6;	DB 10;	Length 608;
Best Local Similarity	81.9%;	Pred. No. 1.5e+100;		
Matches 465; Conservative	0;	Mismatches 97;	Indels 6;	Gaps 1

QY	1	AVPTTTGGCCAGATNTCTGATATTCACATTTTGGCTGTTAAAGGCTCTGGTGGGAAAGACA	60
Db	47	ATGTTTGGCCAGATNTCTGATATTCACATTTTGGCTGTTAAAGGCTCTGGGCGGAAAGACA	106
QY	61	GGCTTGGGTTTCCCTCGATAGTCTTTATTGSAACCCGAGCCAGATCGSACCCCTTAC	120
Db	107	GGCTTGGGGCTTCCCTCTGTCTGTCTTTTTCACAGAGCCGAGACCAATTGGACTCTGTAC	166
QY	121	CATATAGATTAACGTCACGAGATTTATATTCAGSAGAGAAACACACTGACACTTACC	180
Db	167	CATATAGATTAACGACAGATTAATATTAACCGAGAGAGAACGCTCACACTTCACCTGC	226
QY	181	CATGTCGAGATTCCTGACTGCTACTGACTGAGGACCAAGCTGATCCATGCCAGCCCAAC	240

Db	227	TATTTGCACAACCCCTGTACGTGACTACCAAGAGA-----TGACCCTATGGCTTCCAA	280
QY	241	CCCTGTGAACACGGGTGGAGCATGCTCTCCATGGAGACCTTCATATGCAGTGGCCG	300
Db	281	CCCTTACACATGTGTGGGACATGCTCTCAGTGGAGCACCTTCACATGCGCGTGGCG	340
QY	301	GCACCTTCTCTGGGAATAAGTGTGAGAAATGACAAATACGTGCAAGACCAACCCATGT	360
Db	341	GACCTTTCTCTGSAAMAACAGTGTGAGATGTGSAAMAACAGTGTGSAAMAACACCCCTGT	400
QY	361	GGCCGGGAGCCATGTCATTTACCCAGAGTCCCTCTACTACCGCTGTGTGTAAACAC	420
Db	401	GGCCGGGAGAGCTGTCTCATTACTACAGAGTCCCTTTACCAACCGCTGTGCTGAAACAC	460
QY	421	CCTTACACAGGTGCCAAGTGTCTCCCAAGTGGTCTGTATGACAGGCCAAACCCCTGGAG	480
Db	461	CCTTACAGAGGTCTCAGACCTTCACAGAGTGGTCTGTGTGACAGGCCAAATCCCTGMCAA	520
QY	481	AATGGGGGTACTCTCTCCGCGATTAAGGGAGATTCACAGTTCACCTGTGCTGCCGAC	540
Db	521	AAGGTGGACACTCTCTCCGCGACAGAGGGAGAGTTCACAGTTCACCTGTGCTGCCGAC	580
QY	541	CAGTTTCAGAGGGAAATTTCTGTGAATAG	568
Db	581	CAGTTTCAGAGGGAAATTTCTGTGAATAG	608

RESULT	11
Locus	Bf78097.1
DEFINITION	Bf78097.1 902 bp mRNA linear EST J2-UN-200160210549.F01 NC1-CGAP_Kid14 Mus musculus cDNA clone IMAGE:4223710 5', mRNA sequence.
ACCESSION	Bf78097.1
VERSION	Bf78097.1.1 GI:12086004
KEYWORDS	EST,
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE	1 (bases 1 to 902)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: Mac clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LIA9812 row: 1 column: 23
High quality sequence stop: 117.

```

FEATURES
  source
    location/Qualifiers
      1..902
      /organism="Mus musculus"
      /strain="FVB/N"
      /db_xref="taxon:10090"
      /clone="IMAGE:4223710"
      /clone_1lb="NCI_CGAP_K1d14"
      /lab_note="DH10B (TI phage-resistant)"
      /mole="Organ: Kidney; Vector: pGMW-SpOR6; Site_1: NCI;
      Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
      Average insert size 1.75 kb. Constructed by life
      Technologies. Note: this is a NCI_CGAP Library. 1"
BASE COUNT
  223 a 264 c 254 g 161 t
ORIGIN

Query Match      23.2%   Score 390.4;   DB 12;   Length 902;
Best local Similarity 73.1%   Pred. No. 7.7e-100;
Matches 532; Conservative 0; Mismatches 186; Indels 10; Gaps 2;

```

OY	1	AGTTTGGACAGATGTCATGATCCCATGTCTCCTGTAATGCCTTGATGCTCGTGTGGAAACACA	60
Db	52	ATAATTTCAGGAATGTGGTGTCCGTGCTGCTGCTAATGCCCTGTGTGGGAAGTCA	111
OY	61	GCTGTGGGTCCTCCCGATGTCCTTATTTGGAAAGCCGTGACCCAGACTGGACCCCTGCAG	120
Db	112	GTGATTTGGGGCTCTACATGATGTCCTTCATTTGCGCCCCAGACCAGATTTGAGACCCCAGAT	171
OY	121	CAGTATGATTACACTACAGAGATTATATACAGAAGAACACACAGTAGACACTTACC	180
Db	172	GACTATTTACTACACTATAGACAGTCCAGCCCAAGACGAAACCCACGTGTACCCCAACAC	231
OY	181	CATGCTGAANAATCTGACTGTACTACACTAGAGACCAGCTGATCCATGCGACGCCAAC	240
Db	232	ACCCCTGGAAGAACCCCGACTGTGACT-----ATGAAACATGATTCATTCGCAAGTCCAC	285
OY	241	CCCTGTGAACACGGTGGGAGACTGCGCTGCTCATGGGAGACCTTCACATGAGCTGCTG	300
Db	286	CCCTGTGAACACGGCGGGGACTGTATCATACAGAGGGGATACCTTCAGTGTGACGCTGCCA	345
OY	301	GCCTCTTCTGTGGGAATAGTGTGCAAGAAGTGCAAAATACGTGCACAGACACCCATGT	360
Db	346	GGCCCCCTCTGTGGGAGACCGGTSCAGACTGCACAAAAAGTGCAGAAAGCACACCATGT	405
OY	361	GGCGGGGCGCAATGTCATATCCAGAGTCCCTCCACTACACGGCTGTGTCTGTAAACAC	420
Db	406	GTCCATGTGATTTGCTTCATTTACCCAGAAACACCCCTACTACACGGCTGTGTCTGTAAATAC	465
OY	421	CCTTACACAGTCCCGACCTGCTCCCAAGTGSCTCTGTATGCAGGSCCAAACCCCTGCCAG	480
Db	466	CCTTACACAGGACACAGACTGCTCCAAGTGTCTCCGCAATGACAGGCAAAACCCCTGCCAG	525
OY	481	AATGGGGCTACCTGCTCCCGCATTAAGCGGAGATCCAAAGTTCACCTGTGCCCTGTGCCGAC	540
Db	526	AATGGGCGAGTCTGCTCCCGACACAGAGGGAGATCCAGGTTTTACCTGTGCCCTGTGCCGAC	585
OY	541	CAGTTTCAAGGGGAATTTCTGTCAATAGTTCCTATACAGTGCATATTTGCGATGGGCTAC	600
Db	586	CAGTAAATAGGGGAATTTCTGTGAATAGGTCCGGACAGCTGTTATGTGCGTATGGCTAC	645
OY	601	TCTTACCGGGAANAATGATPAGACAGTCAACACGATCGTCCTTACTGAGAACATCC	660
Db	646	TCTTACCGGGAAGTACTAGACAGTCCACAGAACCATGCTTTTTCGGGACTGCCA	705
OY	661	CAGCTCTCTTTGCAGAGAAATTACACATGTTTATGAGAGATGCTTAACCCATGGCATTT	720
Db	706	CTCTCTC----TGGCGGGAGATTCTAGCGCTGTATGCGAGGTCGGAGACCCCGGGAAGC	761
OY	721	GGGGAACA 728	
Db	762	CGGACACA 769	
RESULT 12			
LOCUS	B1220028		
DEFINITION	B1220028	682 bp	mRNA linear EST J1-JUL-2001
ACCESSION	60293491.F1	NCL_CGAP_L19	Mus musculus cDNA clone IMAGE:5098442 5'
VERSION	B1220028		
KEYWORDS	B1220028.1	GI:14673472	
SOURCE	EST.		
ORGANISM	house mouse,		
	Mus musculus		
	Eukaryota; Metazoa;		
	Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;		
REFERENCE	1 (bases 1 to 682)		
AUTHORS	NIH-MGC (http://mgc.ncl.nih.gov/).		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-femail.nih.gov		
	Tissue Procurement: Jeffrey E. Green, M.D.		

FEATURES		source	
		1. 682	
		/organism="Mus musculus"	
		/strain="FVB/N"	
		/db_xref="taxon:10090"	
		/clone="IMAGE:5098442"	
		/clone_lib="NCI_CGAP_L19"	
		/lab_host="DH10B (TI phage-resistant)"	
		/note="Organ: liver; Vector: pCMV-Sport6; Site:1; Notif: Site:2; Salt: Cloned unidirectionally. Primer: oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies Note: this is a NCI_CGAP Library."	
BASE COUNT		166 a	206 c
ORIGIN		169 g	141 t
Query Match	22.9%: Score 386; DB 13; Length 682;		
Best Local Similarity	77.0%: Pred. No. 1.2e-98;		
Matches 511; Conservative	0; Mismatches 145; Indels 8; Gaps 3;		
1	ATGTTGCCAGATGTCGATCCATGCTGCTGCTGTTAATGGCTCTGGTGGAACACA	60	
24	ATATTGTCAGAGATGTTGTGTCGTCGTCGTCGTCGTCATTCGCTCGTGGGAAATCA	83	
61	GGCTTGGGGTCTCCCGATGCTGCTTTATTTGGAAAGCCTGGACCCAGACAGCCCGAC	120	
84	GTCATTTGGGCTCTACATGATGCTCTTCATTCGCGCCCCAGACCCAGATTTGGACCCGAT	143	
121	CAGTATGATTACACTACGAGATTATANTACAGAGAGAACACAGTACACACTTAC	180	
144	GACTATTTCACACTATGAGCGATCCAGCCACAGACACACCCAGCTGTACAGCAGACC	203	
181	CATGTCGAGAAATCTCTGACTGATCACTACGATGAGAGACCAAGCTGATCCATGGACCCAC	240	
204	ACCCCTGAGAACCCCGACTGTACT-----ATGAAAGACGATGATTCATGCGAGTCCAC	257	
241	CCCTGTGAACACGGTGGGACTCCGTCGTCATGAGAGACACTTCATCAGCTGCGCTG	300	
258	CCCTGTGAACACGGGCGGAGCTGTATCATCAGAGGGGATTCCTTGAGTTGCGAGCTGCCA	317	
301	GCCTCTTCTCTGGGAATTAGTGTGAGAAATCGTGAAGAACAGCAACCCATGT	360	
318	GGCCCTTCTGGGAGACCGGTGCCAGATGCAACAAACAGACAGACGACCACTGT	377	
361	GGCGGGGCCAATGTCTCATATCCAGAGAGCTCCCTCACTACGCTGTGTCTGTAACAC	420	
378	GTCCTAGTGATGTCCTCATATCCAGAGACCCCTCACTACGCTGTGCTCAATAATAC	437	
421	CCTTACAGAGTCCACACTCTCTCCCAAGTGGTCTCTTATGCAAGGCAACCCCTGCCAG	480	
438	CCTTACAGGAGACCACTGCTCAAAAGTCTTCGGCATGCGAGGCC--AACCCCTGCCAG	496	
481	AATGGGCTACTCTCTCCGCAATGAGAGATCCAAAGTTCACGTCGCTGCCGAC	540	
497	AATGGGAGTCTGTTCGGACACAGACGAGATCCAGGTTTACGTGTGCTGTCAGAC	556	
541	CAGTTCAAGGGAAATCTGTGAATATAGTTCGTGATACGTCATATGTGTGGATGGCTAC	600	
557	CAGTATATAGGGAAATCTGTGAATATAGTTCGGAGACAGCTGTATGTGTGGATGGCTAC	616	
601	TCTTACCGAGGAAATGAATAGGACAGTCACACAGCATGCGTGCCTTTACTGGAATC	659	
617	TCTTACCGAGGCAAGTGAAGTGAAGCAGTCACACAGCAACCCATGCTTTACTGGAATC	676	
660	CGAC	663	

Db 677 CCAC 680

RESULT 13

LOCUS

DEFINITION 602112402P1 NCI_CGAP_Ki14 Mus musculus cDNA clone IMAGE:4240465

5', mRNA sequence.

ACCESSION BF785781.1 GI:12090817

VERSION

KEYWORDS

SOURCE

ORGANISM house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE NIH-MGC http://mgs.nci.nih.gov/

AUTHORS 1 (bases 1 to 929)

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaaps-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM9855 row: d column: 02

High quality sequence stop: 625.

Location/Qualifiers

1..929

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_image="IMAGE:4240465"

/clone_lib="NCI_CGAP_Ki14"

/lab_host="DH10B (TI phage-resistant)"

/note="Organ: kidney; Vector: pCMV-Sport6; Site: 1; NotI;

Site: 2; SalI; cloned unidirectionally. Primer: oligo dt.

Average insert size 1.75 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library. 1"

BASE COUNT 238 a 269 c 237 g 185 t

ORIGIN

Query Match 22.9%; Score 384.6; DB 12; Length 929;

Best Local Similarity 79.8%; Pred. No. 3.5e-98;

Matches 478; Conservative 0; Mismatches 119; Indels 2; Gaps 2;

372 ATGTCCTATTACCGAGTCTCTCTACTACGAGTGTGTGTAAACACCTTACACAGG 431

19 ATGCTCATTTACCGAAGACACCCCTACTACGCTGTCTCTGCAATACCTTACACGGG 78

432 TCCGACCTCTCCCAAGTGTCTGTATGAGAGGCAACCCCTGCGAGATGGGCTTAC 491

79 ACCAGACTCTCCAAAGTGTCTGTATGAGAGGCAACCCCTGCGAGATGGGCTTAC 138

492 CTGCTCCGCGATTAAGCGAGATTCACCTGTCTCTCTGCGACCAAGTTCAAGGG 551

139 CTGCTCCGCGATTAAGCGAGATTCACCTGTCTCTCTGCGACCAAGTTCAAGGG 198

552 GAAATTCGTGAATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGATGGGCTTAC 611

199 GAAATTCGTGAATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGATGGGCTTAC 258

612 GAAATTAAGTGAAGATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGATGGGCT 671

259 CAAGAGTGAATTAAGTGAAGATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAG 318

672 GCAGGGAATTAAGTGAAGATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGAT 731

319 GCAGGGAATTAAGTGAAGATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGAT 788

QY 732 TTTCGCAAGAACCCAGATGCGGACGAAAGCCCTGCTTATTATTAAGTTACCATTGA 791

Db 379 CTTCGCAAGAACCCAGATGCGGACGAAAGCCCTGCTTATTATTAAGTTACCATTGA 438

QY 792 CAAGGTGAATTAAGTGAAGATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGAT 851

Db 439 GAAAGTGAATTAAGTGAAGATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGAT 498

QY 852 AGAGGAAGCCCTGCTGAGTGTATGAGAGGCAACCCCTGCGAGATGGGCTTACGAGAAAG 911

Db 499 AGTGAAGCCCTTCTGAGTGTATGAGAGGCAACCCCTGCGAGATGGGCTTACGAGAAAG 557

QY 912 TGAGATACGAGAGGAGATTAAGTGAAGATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGAAAG 970

Db 558 CGAGTGAATTAAGTGAAGATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGAAAG 615

RESULT 14

LOCUS BF384535

DEFINITION 602046804F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4196376 5',

mRNA sequence.

ACCESSION BF384535

VERSION BF384535.1 GI:11365840

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE NIH-MGC http://mgs.nci.nih.gov/

AUTHORS 1 (bases 1 to 969)

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaaps-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM533 row: g column: 01

High quality sequence stop: 692.

Location/Qualifiers

1..969

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_image="IMAGE:4196376"

/clone_lib="NCI_CGAP_L19"

/lab_host="DH10B (TI phage-resistant)"

/note="Organ: liver; Vector: pCMV-Sport6; Site: 1; NotI;

Site: 2; SalI; cloned unidirectionally. Primer: oligo dt.

Average insert size 1.9 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library. 1"

BASE COUNT 238 a 278 c 260 g 172 t

ORIGIN

Query Match 22.1%; Score 372.4; DB 12; Length 969;

Best Local Similarity 76.5%; Pred. No. 1.1e-94;

Matches 484; Conservative 0; Mismatches 142; Indels 7; Gaps 2;

QY 1 ATGTTGGCAGAGTGTCTGTATGAGAGGCAACCCCTGCGAGATGGGCTTACGAGAAAG 60

Db 64 ATATTTTGACAGATGTGTGTGTATGAGAGGCAACCCCTGCGAGATGGGCTTACGAGAAAG 123

QY 61 GCCTGGTGTCTCTCTGTATGAGAGGCAACCCCTGCGAGATGGGCTTACGAGAAAG 120

Db 124 GTATTTGGTGTCTCTCTGTATGAGAGGCAACCCCTGCGAGATGGGCTTACGAGAAAG 183

QY 121 CACTATGATTAAGTGAAGATTAAGTGTCTGTATGAGAGGCAACCCCTGCGAGATGGGCTTAC 180

Db	184	GACTTATCTACACCTTATGTGAGCGTCCAGGCCACGACCAAGACCCACGTGTCCACCAAGC	243
Oy	181	CATGCTGGAATTCCTGACTGATCTACATCACTGAGGACCAAGCTGATCTATGTCAGGCCAGC	240
Db	244	ACCCCTGAGAAACCCGCACTGTACT-----ATTAAGACGATGATTCATGCAATCCAGC	297
Oy	241	CCCTGTGAACAGGATGGGAGACGCGCTCGTCGCAATGGGAGACACCTTACATGTGAGTGGCTTG	300
Db	298	CCCTGTGAACACGGCGGGGAGCTTATCATCAGAGGGATACCTCATGATTGAGCTGGCCCA	357
Oy	301	GCTCTCTTCTCTGGGAATAAAGTGTCAAGAAATGCAAAATACGTGCAAGAGCAACCCATGT	360
Db	358	GGCCCTTCTGGGGAAGCGGTGCCAGACTGCACAAACAAAGGGAAGAACACCCATGT	417
Oy	361	GAGCGGGGCCAATCTTCATTTACCCAGAGTCTCTCCCTACTACGCTGTGTCTGTAAACAC	420
Db	418	GTCTATGTGTGATTCCTCATATACCAGAAAGACCCCTACTACGCTGTGCTGCACAAATAC	477
Oy	421	CGTTACAGAGATCCACACTGCTCCCAAGTGTTCTCTATGCGAGGCGCAAAACCCCTGGCAG	480
Db	478	CTTTAAGAGGAGACCACTGCTCCAAAGTGTTCCGCGATGAGGCGCAAAACCCCTGGCAG	537
Oy	481	AATGGGGGACTCTCTGCGGGCATTAAGGGAGATCCAAAGTTCACCGCTGTGCTGCCGAC	540
Db	538	AATGGGGAGTCTGTTCGCCACACAGAGGAGATCCAGGTTTACCTGTGCTGTGCCAGAC	597
Oy	541	CAGTTCAAGGGAATTCCTGTGAATAGTTCCTATACACTGTATGTGGCGATGGCTAC	600
Db	598	CAGTATAAGGGAATTCCTGTGAATAGGTCGGAGACACGTATATGTCCGTATGGCTAC	657
Oy	601	TCTTACCGAGGAATGAATAGACAGTCAAC	633
Db	658	TCTTACCGA-GGCAAGTGAGTAAGACGTCAAC	689

RESULT 15
LOCUS AM475402
DEFINITION
ACCESSION AM475402
VERSION
KEYWORDS
SOURCE
ORGANISM
TITLE
JOURNAL
COMMENT

554 bp mRNA linear EST 24-FEB-2000
un5607.y1 Sugano mouse kidney mki Mus musculus cDNA clone
IMAGE:2505549.5, similar to TR:014520 Q14520 HGF ACTIVATOR LIKE
PROTEIN ;, mRNA sequence.
AM475402
AM475402.1 GI:7045508
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 654)
Marta, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
B., Waller, T., Gibbons, M., Pape, D., Harvey, N., Schurr, R., Ritter
E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Maria M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1600
Fax: 314 286 1810
Email: mousestevatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
WGI:1018401
Seq primer: custom primer used
High quality sequence stop: 482.
Location/Qualifiers
1..654

```
/organism="Mus musculus"  
/strain="C57BL"  
/db_xref="taxon:10090"
```

```

/cclone="IMAGE:2503549"
/cclone_lib="Sugano mouse kidney mkia"
/sex="female"
/dev_stage="adult"
/lib_host="DH10B"
/note="Organ: Kidney; Vector: pME18S-FL3; Site_1: DraIII
(CACCATGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTCGCCCTTTTCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TTCTGGCCCTATCGC), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACCATGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCCTCTCTTAAAGCCGCG and 3' end
primer GCACCTGAGGCTCAGCAACA."
156 a 193 c 174 g 131 t

```

Query Match	22.0%:	Score 370.6:	DB 10:	Length 654:
Best Local Similarity	74.5%:	Pred. No. 2,8e-94:		
Matches 482:	Conservative 0:	Mismatches 159:	Indels 6:	Gaps 1
Qy 1	ANGTTGGCCAGATGTCTGATCTCATGTTCTGCTGTAATAGSGCTCTGSGGAAGACA	60		
Db 13	AAATTTGACAGATGTTGTTGTTCCGGTCCGGTCCGCTTTATGTCCTGAGGGAAGTCA	72		
Qy 61	GGCTTGAGGTTCTCCCTGATGTCTTATTGGAAGACCTGAGACCAGACTGACCCCTGAC	120		
Db 73	GTCAATGTGGCTTCACACTGATGTCCTTCATTTGGGCCCCACACAGATTTGACACCCGAT	132		
Qy 121	CAGTATGATTACAGCTACGACAGATTAATTCAGAAAGAGAACCCAGTACCACTTACC	180		
Db 133	GACATTTACACAGCTATGACATCCACCCAGCAAGAACCCACGATGACGACAGCC	192		
Qy 181	CATGTGAGATCTGACAGCTACTACCTAGCGACCAACCTGATTCATGTCGACGCCAAC	240		
Db 193	ACCCTTGAGACCCCGAGCTGTAAT-----ATGAAAGCATGATTCATGTGCAAGGCCAAC	246		
Qy 241	CCCTGTGAACAGGTGAGGACTGCTGTCGCCATGSGAGACCTGTACATATGACATGCTCG	300		
Db 247	CCCTGTGACACGCGGGAGACTGATCATCAGAGGGATACCTTCAGTTGACAGTCCCA	306		
Qy 301	GCTCTCTTCTCTGGGAATTAAGTCAAGAAAGTCCAAATACGTGCAGAGACACCATGT	360		
Db 307	GGCCCTTCTCGGGGAGACCGGTGTGCANACTGTGCANAGCAAGGCAAGCAACACCATGT	366		
Qy 361	GGCCGGGGCCAAATGTCATTAATACCAGAGTCTCCCTACACGCTGTTGCTTAACAC	420		
Db 367	GTCCATGATGATTTGCTTCATTACCCAAAGACACCCCTACACGCTGTGCTGACATAC	426		
Qy 421	CCCTACAGAGTCCAGCGCTGCCAAGTGGTTCCTCATGCGAGGCAAAACCCCTGCGAG	480		
Db 427	CCTTACACGAGACACAGACTGCTCCAAAGTGTGGGATATGACGAGCAAAACCCCTGCGAG	486		
Qy 481	AATGGGATACACGTGCCCGGCAATAGAGGGAATCAAGTTCAACCTGTGCTGCGGAC	540		
Db 487	AATGGGAGATCTGTGTTCCCGACACAGACGGAATTCAGAGGTTACCTGTGCTGTCCGAC	546		
Qy 541	CAGTTCAAGGGAATTTCTTGGAATTAAGTTCTGATGACCTGATGTTGGGATGGCTAC	600		
Db 547	CAGGTATTAAGGGGAATTTCTGTGACATTAAGTGGGAGCGATGTGATTCGGTGATGGCTAC	606		
Qy 601	TCTTACGAGGGAAATTAATAGACAGTCAACACAGCATGCGTCTG 647			
Db 607	TCTTACGAGCCAAAGGAGTGAACAGGCAAGCGAAGCGAATGCT 653			

Search completed: March 6, 2003, 19:37:28
Job time : 2189.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 15:27:01 ; Search time 65.5 Seconds
(without alignments)
7879.954 Million cell updates/sec

Title: US-09-912-559-1

Perfect score: 1683

Sequence: 1 atgttgcacgaagtctctga.....aaagtgaagtgcttctaa 1683

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/lna/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/lna/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/lna/PCrus_COMB.seq:*

6: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117.6	7.0	2033	1 US-08-148-910-14	Sequence 14, Appl
2	117.6	7.0	2033	1 US-08-448-937A-14	Sequence 14, Appl
3	96.2	5.7	970	1 US-08-148-910-3	Sequence 3, Appl
4	96.2	5.7	970	1 US-08-448-937A-3	Sequence 3, Appl
5	75.4	4.5	1065	1 US-08-427-640-1	Sequence 1, Appl
6	75.4	4.5	1065	1 US-08-427-640-5	Sequence 5, Appl
7	75.4	4.5	1068	1 US-08-427-640-3	Sequence 3, Appl
8	75.4	4.5	1137	4 US-09-553-498-9	Sequence 9, Appl
9	75.4	4.5	1137	4 US-09-618-869-9	Sequence 9, Appl
10	75.4	4.5	1314	2 US-08-811-949-48	Sequence 48, Appl
11	75.4	4.5	1955	2 US-08-883-795A-39	Sequence 39, Appl
12	75.4	4.5	2457	6 5344773-1	Patent No. 5344773
13	75.4	4.5	7360	1 US-08-286-740-1	Sequence 1, Appl
14	75.4	4.5	7360	5 PCT-US95-09576-1	Sequence 1, Appl
15	73.8	4.4	329	1 US-08-148-910-2	Sequence 2, Appl
16	73.8	4.4	329	1 US-08-148-910-13	Sequence 13, Appl
17	73.8	4.4	329	1 US-08-148-910-13	Sequence 13, Appl
18	73.8	4.4	329	1 US-08-448-937A-2	Sequence 13, Appl
19	73.8	4.4	329	1 US-08-448-937A-13	Sequence 13, Appl
20	73.8	4.4	1065	2 US-08-811-949-60	Sequence 60, Appl
21	73.8	4.4	1068	1 US-08-137-116-2	Sequence 2, Appl
22	73.8	4.4	1068	1 US-08-427-640-7	Sequence 7, Appl
23	73.8	4.4	1068	2 US-08-811-949-44	Sequence 44, Appl
24	73.8	4.4	1068	2 US-08-811-949-46	Sequence 46, Appl
25	73.8	4.4	1068	2 US-08-811-949-52	Sequence 52, Appl
26	73.8	4.4	1068	2 US-08-811-949-58	Sequence 58, Appl
27	73.8	4.4	1163	6 523256-3	Patent No. 523256
				2 US-08-558-269-5	Sequence 5, Appl

28	73.8	4.4	1163	4 US-09-810-882-5	Sequence 5, Appl
29	73.8	4.4	1170	2 US-08-811-949-64	Sequence 64, Appl
30	73.8	4.4	1170	2 US-08-811-949-66	Sequence 66, Appl
31	73.8	4.4	1114	2 US-08-811-949-50	Sequence 50, Appl
32	73.8	4.4	1114	2 US-08-811-949-54	Sequence 54, Appl
33	73.8	4.4	1114	2 US-08-811-949-56	Sequence 56, Appl
34	73.8	4.4	1419	2 US-08-811-949-62	Sequence 62, Appl
35	73.8	4.4	1738	6 5200340-1	Patent No. 5200340
36	73.8	4.4	1848	3 US-08-814-412-10	Sequence 10, Appl
37	73.8	4.4	1974	2 US-08-811-949-38	Sequence 38, Appl
38	73.8	4.4	2101	2 US-08-811-949-42	Sequence 42, Appl
39	73.8	4.4	2162	1 US-08-119-512-3	Sequence 3, Appl
40	73.8	4.4	2162	1 US-08-488-0158-3	Sequence 54, Appl
41	73.8	4.4	2162	1 US-08-488-0158-25	Sequence 25, Appl
42	67.8	4.0	1724	6 5200340-5	Patent No. 5200340
43	67.8	4.0	2497	6 5185259-2	Patent No. 5185259
44	66.4	3.9	2544	4 US-09-518-046-3	Sequence 3, Appl
45	66	3.9	1233	1 US-08-254-922-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-148-910-14
Sequence 14, Application US/08148910
Patent No. 5465593
GENERAL INFORMATION:
APPLICANT: Takeshi SHIMOMURA et al.
TITLE OF INVENTION: No. 5465593el Protein and Gene Encoding Said Protein
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch,
MEDIUM TYPE: 500 KB Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Motoprefect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/148,910
FILING DATE: No. 5465593ember 5, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ. ID NO. 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE: human
ORGANISM: human
IMMEDIATE SOURCE:
LIBRARY: Pre-made Lambda phage library,
LIBRARY: human liver(49, male) cDNA Library (Stratagene)
US-08-148-910-14

Query Match	7.0%	Score 117.6:	DB 1	Length 2033:
Best Local Similarity	49.8%	Pred. No. 6.8e-25:		
Matches 607:	Conservative 0:	Mismatches 559:	Indels 54:	Gaps 10
OY	469	AACCCCTGCCAAGATGGGGCTACCTGCTCCGGCCATAAGCGGAGATCCAAAGTTCACCTGT	528	
Db	742	AGCCCTTGCTTGAAGACGGGGACCTCCACCTGATACGAGGCCACCGGGACACCTGTGT	801	
OY	529	GGCTGCCCGACACAGTTTCAGGGGAATTCGTWGAATAGGTTCTATATA---CTGCAT	585	
Db	802	GCCTGCCACCAAGGCTTCGCTGACGCGCTCTGCAACATCAGCCTATATAGCGCTGCTTC	861	
OY	586	GTTGGCAGTGCCTACTCTTACCGAGGGAATGAATVAGACAGTCAACCAAGCATCCGTGC	645	
Db	862	TTCGGGAGAGGAGCTGGGTACCGCTGGCCGTCGACACACCTCAAGCTTCGGGCTCAGCTGC	921	
OY	646	CTTTACTGGAATCCCACTCTCTTTCGACGGAATTAACAACATGTTTAVGAGATATCT	705	
Db	922	CTGGCCTGGAATCCGATCTGCTTACGACGATCGACGATGATCCTGGGCGCCCGC	981	
OY	706	GAACCCATGAGATTTGGGGAACAAATTTCTGCAGAAACCCAGATCGAGCAAAAGCC	765	
Db	982	GGCCTCTGGGCGCTGGGCCCCCATGGCTCTCTCCGGAATCCGGACATGACCAAGAGGCC	1041	
OY	766	TGGTGGTTTATTAAGTTACCAATGACAGGAATGGGATACGTATGATCTGCACGC	825	
Db	1042	TGGTGTAGCT---GGTAAAGAACAGCGCGCTCTCTGGAGTATCGCGCTGGAGAGC	1098	
OY	826	TGCTCAGCCAGAGACGTTGCCTTACCAGAGAAAGCCCACTAGACCTATCAACAGCTT	885	
Db	1099	TGGGAATCCCT-----CACAGAGTCACATGTCACCGGATCTCCTGGCGACCTTG	1149	
OY	886	CCGGGGTTGACTCTGTGGAAAGACTGATAGACAGAGAAATCAAGAAATCTAT	945	
Db	1150	CTGAGACCGAGCTCCCGGGGCGCCAGGCGCTCGGACAGAGACCAAGAGAGACGTTTC	1209	
OY	946	GGAGGTTTAAAGAGACGGGGGGGACAGCAACCATGGAGGGCTCCCTCCAGTCTCGGTG	1005	
Db	1210	CTGGGCGCACGATATCATGGGGGCTCTCTCTGCTGCCGGCTCGACCCCTG-----	1262	
OY	1006	CCCTGACCATCTCCATGCCCCAGGGGACATTTCTGTGGTGGGGCGGTGATTCACCCCTGC	1065	
Db	1263	--CGTGGCGGATCTACATCGGGGGACACTTCTCGCGGGGAGACCTGGTCCACACCTGC	1320	
OY	1066	TGGGTCTCATCGCTGGCCACTGACCGACATTAATAAACACAGACATCTA-----AGGTG	1119	
Db	1321	TGGGTGTGTCTGGCGGCCCACTCTTCTCCACAGCCCCCCAGAGACAGCTCTCGGTG	1380	
OY	1120	GTCCTAGGGGACCAAGACCTGAAGAAGAATTCATVAGCAAGACGTTTAGGCTGGAG	1179	
Db	1381	GTCCTGGGCGACACTTCTTCAACCCGACAGAGGACGTGACGACGACCTTCGCAATCCAG	1440	
OY	1180	AAGATATTCAAGTACAGCCACTCATATGAAAGATGAGTATGCCCAATATATTTCA	1239	
Db	1441	AAGTATATCCGTATACACCTGTACTGTGGTTCAAACCCAGCAGCAGC---GACCTGCTC	1497	
OY	1240	TTCCTCAAGTTAAAGCCAGTGAATGTCTACTGTCTCTAGATTCAAATATCTGAAGCT	1299	
Db	1498	CTGATCCGGCTGAAGAAAGAAAGGGGACCGCTGTGCACACGCTCGGACGTTCTGCAAGCC	1557	
OY	1300	GTTGCTGCTCCCTGAT-----GGTCCCTTTCCTCTGGGAGTGAAGTGGACATCTCTGGC	1353	
Db	1558	ATCTGCTTCCCGAGGCCCGGAGACACTTTCGCCGAGBACCAACATGCCAGTATCGGGC	1617	
OY	1354	TGGGG-----TGTTACAGAAACAGGAAAGGGTCCCGCCAGCTCTCGATGCCAAATGC	1407	
Db	1618	TGGGGGCACTTGGATGAGAACGTGAACGGGCTATCTCAACATCTCCTGGGGAGGCCCTGTGTC	1677	
OY	1408	AAGCTTATGGCAACACTTTGTCTCACTCCGCGCAACTCTATACCAACATGATATATAC	1467	
Db	1678	CCCTGTGGCGGACCAAGATGTGACAGACCTCTGAGGTCTTAGGGCCCGACATCAAGCCCC	1737	
OY	1468	AGTATGATCTGTGACGAAATTTTCAGAAACCTGGGCAAGACATCTGCCAGGCTACTCTT	1527	

Dd	1738	AACATCTCTGTGGCGGCGTACTT---	CGACTGCAGATCCGACGCGTCGCCAGGGGAACTCA	1794
Oy	1528	GAAGGCCCCCTGACCTGTGAGAAGACGCGCACC	TACTACTGTAAGGATAGTAGCGTG	1587
Dd	1795	GGGGGGGCCCCCTGGCCCTGGGAGAAAGACGGCG	CTGCGTTTACCTTACGACGCAATCATCAGCTGG	1854
Oy	1588	GGCGCTGGAATTGTGGG-----AAGAGCGCAG	GGGGTCTCACACCCCAATTCACAATTCCTGT	1641
Dd	1855	GGGAGAGCGCTGGGGGCGGCGCTCCACACACGG	CGGGGGGTCTACACCGCGGTGGCCACAACTATGTG	1914
Oy	1642	AATTGATGATCAAAGCCACCATT	1661	
Dd	1915	GACTGGATCMACGACCGGAT	1934	
<p>RESULT 2 US-08-448-937A-14 ; Sequence 14, Application US/08448937A ; Patent No. 5677164 ; GENERAL INFORMATION: ; APPLICANT: Takeshi SHIMOMURA et al. ; TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein ; NUMBER OF SEQUENCES: 14 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: Wenderoth, Lind & Ponack ; STREET: 805 Fifteenth Street, N.W., #700 ; CITY: Washington ; STATE: D.C. ; COUNTRY: U.S.A. ; ZIP: 20005 ; COMPUTER READABLE FORM: ; MEDIUM TYPE: Diskette, 5.25 inch, ; MEDIUM TYPE: 500 Kb Storage ; COMPUTER: IBM Compatible ; OPERATING SYSTEM: MS-DOS ; SOFTWARE: Wordperfect ; CURRENT APPLICATION DATA: ; APPLICATION NUMBER: US/08/448,937A ; FILING DATE: May 24, 1995 ; CLASSIFICATION: 435 ; PRIOR APPLICATION DATA: ; APPLICATION NUMBER: 08/148,910 ; FILING DATE: No. 5677164ember 5, 1993 ; ATTORNEY/AGENT INFORMATION: ; NAME: Warren M. Cheek, Jr. ; REGISTRATION NUMBER: 33,367 ; REFERENCE/DOCKET NUMBER: ; TELECOMMUNICATION INFORMATION: ; TELEPHONE: 202-371-8850 ; TELEFAX: 202-371-8856 ; TELEX: ; INFORMATION FOR SEQ ID NO: 14: ; SEQUENCE CHARACTERISTICS: ; LENGTH: 2033 base pairs ; TYPE: nucleic acid ; STRANDEDNESS: double ; TOPOLOGY: linear ; MOLECULE TYPE: cDNA ; ORGANISM: human ; IMMEDIATE SOURCE: ; LIBRARY: pre-made lambda phage library, ; LIBRARY: human liver(49, male) cDNA library (Stratagene) US-08-448-937A-14</p>				
Oy	469	AACCCCTGGCAACAATGGGGGCTACCTCTCCCGCATAAGCGAGATCCAAATTCACCTGT	528	
Dd	742	AGCCCTGCTCTTAAGGGGGGACCTGTGCACACTGATCGTGGCCACCGGAGACCAACCGTGT	801	
<p>Query Match 7.0%; Score 117.6; DB 1; Length 2033; Best local Similarity 49.8%; Pred. NO. 6.8e-25; Matches 607; Conservative 0; Mismatches 559; Indels 54; Gaps 10</p>				


```

Db      321 CTGGCCAGCACTTCTTCAACCGCAGCAGCGAGCGTAGCGCAGCCTTCGGCATCGAAG 380
OY      1183 ATATTCAAGTACACCCACATCATATGAAGATGATATCCCAATGATATATTCATTG 1242
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      381 TACATCCCGTACACCCGCTACTCGGTTCACACCCCA---GGAGCCACAGACCTCGCTCG 437
OY      1243 CTCAGTTAAAGCCATGTGATGTCTGTCTTCAATTCAAATACGTGAAGACTGTG 1302
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      438 ATCCGGCTGAAGAAAGAAAGGACCGCTGTGCACACGCTCGCAGATTCGCGCCATTC 497
OY      1303 TCGTTCCTGATG-----GCTCTTCCCTCTGGAGTGAAGTCCACATCTTGCTGG 1356
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      498 TCGCTGCCGAGCCCGGACACACTTCCCGAGAGCACAAGTCCAGATGTGGCGCTGG 557
OY      1357 GG-----TGTACAGAACAGAAAAAGGTCGCCGACGCTCTGATGTCGAAGTCAG 1410
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      558 GCCCACTTGATGAGAACGTGAGCGGCTACTCCAGCTCCCTGGGAGGCCCTGGTCCC 617
OY      1411 CTGATTGCCAACACTTTGTGCAACTCCCGCAACTCTATCCACATGATGATGACAT 1470
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      618 CTGTGCGCCGACACAGATGACAGCCTTAGAGTCTACGCGCCGACATCAGCCCCAAC 677
OY      1471 ATGATCTGTGACGAATCTTCAGAACTGGGCAAGACACTGCCAGGTGACTGTGA 1530
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      678 ATGCTCTGTGCGGCTACTT---CGACTGCAAGTCCGAGCGCTCGAGGGGACTAGGG 734
OY      1531 GGCCCCCTGACCTGTGAGAAGAGCGCACTACTACGTCTATGGGATAGTGAAGCTGGGCG 1590
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      735 GGGCCCTGCGCTCGAGAGAAAGCGCTGCTACTCTACGATCATCAGCTGGGCT 794
OY      1591 CTGAGATGTGG-----AAGAGGCCAGGGGCTTACACCCCAAGTACCAATTCCTGAAT 1644
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      795 GAGCGCTGGGGCGGCTCCACAAACCGGGGTCTACACCGCGCTGGCCAACTATGTGAC 854
OY      1645 TGGATCAAGCCACCAT 1661
        ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      855 TGGATCAAGACCGGAT 871
        ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-08-448-937A-3
; Sequence 3, Application US/08448937A
; Patent No. 5677164
; GENERAL INFORMATION:
; APPLICANT: Takeshi SHIMOMURA et al.
; TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch,
; MEDIUM TYPE: 500 Kb Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,937A
; FILING DATE: May 24, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/148,910
; FILING DATE: NO. 5677164ember 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mairen M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850

```

```

; TELEFAX: 202-371-8856
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; LIBRARY: Pre-made Lambda phage Library, human liver (49, male)
; LIBRARY: cDNA Library (Stratagene)
US-08-448-937A-3

Query Match      5.7%; Score 96.2; DB 1; Length 970;
Best Local Similarity 53.0%; Pred. No. 1e-18; Mismatches 288; Indels 30; Gaps 6;
Matches 359; Conservative 0;

OY      1009 CTGACATCTCCATGCCCCAGGACCTTCTGTGTGGGCGCTGATCCACCCTGCTGG 1068
        || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      201 CTGGCGGCATCTACATCGGGGACAGCTTCTGGCGCGAGCGCTGTCCACACCTGCTG 260
OY      1069 GTGTCTACTGTGCGCCACTGCGACCGACATTAACCAAGACA-----TCTAAGGTG 1122
        || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      261 GTGTGTGAGCGCCGACACTGCTTCTCCACAGCCCCCGAGGACAGCGCTCCGCTGTG 320
OY      1123 CTAGGAGACAGACACTGAGAAAGAAATTTTCATGACGACGCTTTAGGTGAGAG 1182
        || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      321 CTGGGCGAGACTTCTTCAACCGCAGCAGGAGCTGACGACGACCTTGGCATCGAAG 380
OY      1183 ATATTCAAGTACAGCCACTACATGAAGAAAGATGAGATTCCCAATGATATTCATTG 1242
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      381 TACATCCCGTACACCCGCTACTCGGTGTACACCCA---GGACACAGACCTGTGCTG 437
OY      1243 CTCAGTTAAAGCCAGTGATGCTACTGTGCTGTAGAAATCCAAATACGTAACACTGTG 1302
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      438 ATCCGGCTGAAGAAAGAAAGGAGACCGCTGTGCCACAGCGCTGAGTTGTGCACCCATC 497
OY      1303 TGTGTGCTCATG-----GTCCTTCCCTCTGGAGTGAAGTGCACATCTGTGCTGG 1356
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      498 TCGCTGCCGAGCCCGGACGACCTTCCCGCAGAGCACAAGTGGCAGATTGGCGGCTGG 557
OY      1357 GG-----TGTACAGAACAGAAAAAGGTCGCCGACGCTCTGATGTCGAAGTCAAG 1410
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      558 GCCCACTTGATGAGAACGTGAGCGGCTACTCCAGCTCCCTGGGAGGCCCTGTGTCGCC 617
OY      1411 CTGATTGCCAACACTTTGTGCAACTCCCGCAACTCTATGACACATGATGATGACAGT 1470
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      618 CTGTGCGCCGACACAGATGACAGCCTTAGAGTCTACGCGCCGACATCAGCCCCAAC 677
OY      1471 ATGATCTGTGACGAATCTTCAGAACTGGGCAAGACACTGCCAGGTGACTGTGA 1530
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      678 ATGCTCTGTGCGGCTACTT---CGACTGCAAGTCCGAGCGCTGACCGAGGACTCAAGG 734
OY      1531 GGCCCCCTGACCTGTGAGAAGAGCGCACTACTACTCTATGGGATAGTGAAGCTGGGCG 1590
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      735 GGGCCCTGCGCTCGAGAGAAAGCGCTGCTTACCTTACCGCATCATCGCTGGGCT 794
OY      1591 CTGAGATGTGG-----AAGAGGCCAGGGGCTTACACCCCAAGTACCAATTCCTGAAT 1644
        || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      795 GAGCGCTGGGGCGGCTCCACAAACCGGGGTCTACACCGCGCTGGCCAACTATGTGAC 854
OY      1645 TGGATCAAGCCACCAT 1661
        ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      855 TGGATCAAGACCGGAT 871
        ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-08-427-640-1
; Sequence 1, Application US/08427640
; Patent No. 5658788

```



```
OY 1028 AGGGCCACTTCTGTGGTGGGGCTGATCCACCCCTGCTGGTGTCTCACTGCTGCCACT 1087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 461 AGCGTTCTGTGGGGGGCATATCATCACTCTCTCTGATTTCTCTGCGGCCACT 520
OY 1088 GCACCGA-----CATMAAACACGACATCTAAAGTGTGTAGGGGACCGAGCTCGA 1141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 521 GCTTCCAGAGAGGTTCCGCCCCACGACGTGAGGATCTTGCGGACAGACATACCGG 580
OY 1142 AGAAGAAGATTTCTAGACAGACTTTAGGTGAGAGATATTCAAGTACAGCCACT 1201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 581 TGTTCCTGGCGAGAGAGAGAAATTTGAAGTCGAAAATATCATTTGCCATAGGAT 640
OY 1202 ACATGAAAGATGATGATTTCCCAATGATATTGCTTCAAGTTAAAGCCAGTGG 1261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 641 TCGAT-----GATGACACTTACGACATGACATTTGGGCTGCGACGATGGAATG 694
OY 1262 ATGGTCACTGTCTCTGACAAATACGTAAGACTGTGTCTGCTGCTGATGGTCT 1321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 695 CGTCCCGCTGTGGCCAGAGAGACGCTGTCCGCACTGTGTCTTCCCGCGGAC 754
OY 1322 TTCCCT-----CTGGAGTGTGAGTCCACATCTGTGCTGGGTGTACAGAAA 1369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 755 TGCAGCTGCGGAGCTGACGAGGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCTTGT 814
OY 1370 CAGGAAAAGGTCGCCGACGCTCTGATGCCAAAGTCAAGCTGATTGCCAAGCTTGT 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 815 CTCCTTTCTATTGGAGCGGTGAAGGAGGCTCATGTACACATGTACCATACCGCT 874
OY 1430 GCACACTCCGGCCACTGTATGACCATGATGTATGACATGATGATCTGTGAGAAATC 1489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 875 GCACATACACATTTACTTTACAGACAGTACCGGACATGCTGTGTGTGAGACA 934
OY 1490 TTCAAAACCTG-----GCAAGACACTGCCAGGTGACTGTGAGGCC 1534
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 935 CTCGAGCGGGGCCCCAGGCAACTTGCACGACGCTCCAGGCGCATTTGGAGGCC 994
OY 1535 CCCTGACCTGTGAGAGAGGCGACCTACTACGCTATGGGATAGTGGGCGCTGG 1594
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 995 CCCTGTGTCTGACAGATGCGCGCATGACTTTGGTGGCATCATACGTGGGCGCTGG 1054
OY 1595 AGTGTGGGAAGG-----CCAGGGGTCTACACCCCAAGTTACCAATTTCTGAATTGGA 1648
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1055 GCTGTGACAGAGATGTCGGGTGTGTACACCAAGTTACCAATCTACCTAGCTGGA 1114
OY 1649 TCAAGCCACCAT 1661
      ||| ||| |||
DB 1115 TTCTGTACACAT 1127
      ||| ||| |||

RESULT 9
US-09-618-869-9
; Sequence 9, Application US/09618869
; Patent No. 6455279
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorthée
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwartz, Elisabeth
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND
; TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR
; FILE REFERENCE: 20381
; CURRENT APPLICATION NUMBER: US/09/618, 869
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: EP99114811.5
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
```

```
NAME/KEY: CDS
LOCATION: (1)..(1137)
US-09-618-869-9

Query Match 4.5%; Score 75.4; DB 4; Length 1137;
Best Local Similarity 50.8%; Pred. No. 1.7e-12;
Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

OY 1028 AGGGCCACTTCTGTGGTGGGGCTGATCCACCCCTGCTGGTGTCTCACTGCTGCCACT 1087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 461 AGCGTTCTGTGGGGGGCATATCATCACTCTCTGATTTCTCTGCGGCCACT 520
OY 1088 GCACCGA-----CATMAAACACGACATCTAAAGTGTGTAGGGGACCGAGCTCGA 1141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 521 GCTTCCAGAGAGGTTCCGCCCCACGACGTGAGGATCTTGCGGACAGACATACCGG 580
OY 1142 AGAAGAAGATTTCTAGACAGACTTTAGGTGAGAGATATTCAAGTACAGCCACT 1201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 581 TGTTCCTGGCGAGAGAGAGAAATTTGAAGTCGAAAATATCATTTGCCATAGGAT 640
OY 1202 ACATGAAAGATGATGATTTCCCAATGATATTGCTTCAAGTTAAAGCCAGTGG 1261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 641 TCGAT-----GATGACACTTACGACATGACATTTGGGCTGTGACGTGAAATTCGAT 694
OY 1262 ATGGTCACTGTCTCTGACAAATACGTAAGACTGTGTCTGCTGCTGATGGTCT 1321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 695 CGTCCCGCTGTGGCCAGAGAGACGCTGTCCGCACTGTGTCTTCCCGGCGGAC 754
OY 1322 TTCCCT-----CTGGAGTGTGAGTCCACATCTGTGCTGGGTGTACAGAAA 1369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 755 TGCAGCTGCGGAGCTGACGAGGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCTTGT 814
OY 1370 CAGGAAAAGGTCGCCGACGCTCTGATGCCAAAGTCAAGCTGATTGCCAAGCTTGT 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 815 CTCCTTTCTATTGGAGCGGTGAAGGAGGCTCATGTACACATGTACCATACCGCT 874
OY 1430 GCACACTCCGGCCACTGTATGACCATGATGTATGACATGATGATCTGTGAGAAATC 1489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 875 GCACATACACATTTACTTTACAGACAGTACCGGACATGCTGTGTGTGAGACA 934
OY 1490 TTCAAAACCTG-----GCAAGACACTGCCAGGTGACTGTGAGGCC 1534
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 935 CTCGAGCGGGGCCCCAGGCAACTTGCACGACGCTCCAGGCGCATTTGGAGGCC 994
OY 1535 CCCTGACCTGTGAGAGAGGCGACCTACTACGCTATGGGATAGTGGGCGCTGG 1594
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 995 CCCTGTGTCTGACAGATGCGCGCATGACTTTGGTGGCATCATACGTGGGCGCTGG 1054
OY 1595 AGTGTGGGAAGG-----CCAGGGGTCTACACCCCAAGTTACCAATTTCTGAATTGGA 1648
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1055 GCTGTGACAGAGATGTCGGGTGTGTACACCAAGTTACCAATCTACCTAGCTGGA 1114
OY 1649 TCAAGCCACCAT 1661
      ||| ||| |||
DB 1115 TTCTGTACACAT 1127
      ||| ||| |||

RESULT 10
US-08-811-949-48
; Sequence 48, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIMA, MINO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
```


Fri Mar 7 09:21:09 2003

us-09-912-559-1.rni

Page 9

Oy	1202	ACATGTGAAGATGTATGATATCCCGACAAATGATATGATCTCAAGTTAAAGCCATGTG	1261
Db	1283	TCGAT-----GATACACTTACGACAAATDACAATTCGCTGTCCACCTGAATATCGGAT	1336
Oy	1262	ATGTGTCAGTGTCTTAAATATCAATATACGTGAAGACTGTGTGCTTGCCTGATGTGCTCT	1321
Db	1337	CGTCCGCTGTGCTCCGAGGAGCGAGCGCTGTGCTCCGACATGTGTGCTTCCCGCGGGAGCC	1366
Oy	1322	TTCCCT-----CTGGAGTAGTGGACACATCTCGCTGGGGGTGTATACAGAA	1369
Db	1397	TGCGAGCTGCCGGACACGACGAGGTGTAGCTGCTCCGGCTACGGCAACATGAGGCCCTGT	1456
Oy	1370	CAGAAAAAGGTCCCGCAGCTGCTGGATGCCAAGTGAACCTATATGCCACACTTTGT	1429
Db	1457	CTCCTTTTATTCGAGAGGGGCTTGAAGBAGGCTCATGTAGACTTACCCATCCAGCGCT	1516
Oy	1430	GCAACCTCCGCCCACTCTATGACACATGATGATGACATGATGATGCTGGAGAAATC	1489
Db	1517	GCACTATCAACATTTACTTATTAACAGAAAGTACCAGAACATCTGTGTGCTGGAGACA	1576
Oy	1490	TTGCAGAACCTG-----CCAAAGCACCTGGCAAGGTGATCTGGAGGCC	1574
Db	1577	CTGGAGGCGCGCGGCCAGGCAAACTTGCAGAGACGCTGCGAGGGGATTCGGGAGGCC	1636
Oy	1535	CCCTGACCTGTGAGAAAGACGCGCACTTACTACCTATAGGATATAGAGACTGGGGCTGG	1594
Db	1637	CCCTGTGTGTGTAGAAAGATGGCGCATACCTTTGTGTGGCATATCACTATGGGGCGTGG	1696
Oy	1595	AGGTGTGAAAGG-----CGAGGGCTTACACCCAAAGTTACCAATTCCTGAAATTGA	1648
Db	1597	GCGTGTGACAAAGAAATGTCCCGGATGTATACCAAGTTTACCACCTACTAGACTGA	1756
Oy	1649	TCAAAGCCACCAT	1661
Db	1757	TTGTGTACACAT	1769

RESULT 12
 5344773-1
 Patent No. 5344773
 APPLICANT: NCI, CHA-MER, HSUNG, NANCY REDDY, VERMURI B.,
 LEMONT, JEFFREY F., DACKOWSKI, WILLIAM DOUGLAS, RICHARD;
 COLE, EDWARD S., PURCELL JR., RICHARD D., LAU, DAVID TAI-YOI
 TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN
 ACTIVATOR PRODUCED BY RECOMBIANT DNA
 NUMBER OF SEQUENCES: 6
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/782,686
 FILING DATE: 01-OCT-1985
 Prior APPLICATION DATA:
 APPLICATION NUMBER: 656,770
 FILING DATE: 01-OCT-1984
 SEQ ID NO:1
 LENGTH: 2457
 5344773-1

Query Match	4.58;	Score 75.4;	DB 6;	Length 2457;
Best Local Similarity	50.88;	Score 75.4;	DB 6;	Length 2457;

Statistics	Conservative	Mismatches	Indels	Gaps
Matches	342;	0;	286;	45;
Mismatches	342;	0;	286;	45;
Indels	342;	0;	286;	45;
Gaps	342;	0;	286;	45;

[illegible]

Db	1206	TCGAT-----GATCACACTTACGACAAATACATTTCGCTGCTGACGTGAATTCGGATT	1255
OY	1262	ATGGTCACTGTGCTGTAAATATCCAAATAGCTGAAGCTGTGTGCTTGTGCTGTATGTGCTCT	1321
Db	1260	CGTCCCGCTGTGTGCTCCAGAGACAGCGCTGTCCGCACGTGTGTGCTTGTCCCTCCCGCGGAC	1315
OY	1322	TTCCCT-----CTGGAGTGAATGCCATCTCTGGCTGGGGTGTTTACAGAA	1365
Db	1320	TGCAGCTGCCGAGACTGACAGGAGATGTGAGCTTCCGGCTACGCGCAAGCATAGAGCCCTTGT	1375
OY	1370	CAGAAAGGATCCCGCAGCATCTCTGGATGCCAAAGTCAAGCTATTGGCCAAACTTTGT	1422
Db	1380	CTCCTTTCTATTTCGAGGCGCTGAAGAGAGCTATGTACAGTCTACCCATTCAGCGCT	1435
OY	1430	GCAATCCCGCCCACTCTATGACCACATGATTGATGACAGATGATCTGTGCAGGAATC	1489
Db	1440	GCACATCACAACTTACTTAACTACAGAACATCAACCAACATGCTGTGTGCTGAGAAC	1495
OY	1490	TTGAGAACTCTG-----GCAAGCAACCTGGCAGAGGTGACTGTGAGAGGC	1534
Db	1500	CTGGAGAGCGGCGGCCCAAGCAAACTTGTGAGAGCGCTGCAGAGCGATTCGGAGGCGC	1555
OY	1535	CCCTGACTGTGAGAGAGCAGCGCACTTACAGCTATGGGATATGTAGCTGGAGGCGCTGG	1594
Db	1560	CCCTGTGTGTGTGAACGATGCCGCGATGACTTTGGTGGGCATCATGCTGGGGCGCTGG	1619
OY	1595	AGGTGTGAAGAG-----CAAGGGGTATACACCAGTTTACCAATTCCTGTAATGGA	1648
Db	1620	GGTGTGACAAAGAGATGTCCCGGTGTGTACACCAAGGTATACCACTTACTTATGACTGGA	1679
OY	1649	TCAAAGCCACCAT	1661
Db	1680	TTGTGTACAACT	1692

RESULT 13
US-08-286

; Sequence 1, Application US/08286740
; Patent No. 5561053

GENERAL INFORMATION

APPLICANT: Crowle

TITLE OF INVENTION: ME

NUMBER OF SENTENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San F

STATE: CALIF
COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 in

COMPUTER: IBM PC COM

OPERATING SYSTEM: PC
SOFTWARE: matlab /Com

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/286,740

FILING DATE: 05-AUG-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER:

REFERENCE/DOCKET NUMBER: 798

TELEPHONE: 415/335-1984

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 1:

```

: SEQUENCE CHARACTERISTICS:
:   LENGTH: 7360 bases
:   TYPE: nucleic acid
:   STRANDEDNESS: double
:   TOPOLOGY: linear
US-08-286-740-1

```

```

Query Match          4.5%; Score 75.4; DB 1; Length 7360;
Best Local Similarity 50.8%; Pred. No. 4.4e-12;
Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

```

```

1028 AGGGCCACTTCTGTGTGGGCGCTGATCCACCCCTGGTGGTGTCACTGCTGCCACT 1087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2634 AGGGTTCCTGTGGGGGGGCATATCATAGCTCCCTGGATTCCTCTGCGCCGAC 2693
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1088 GCACCCA-----CATMAAACCGACATCTAAAGTGTGTAGGGGACCAGGACTGA 1141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2694 GCTTCCAGAGAGGTTTCCGCCCCACCTGACGCTGATCTTGGGAGAAACATACCGGG 2753
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1142 AGAAGAAGATTTTCATGACGACGCTTTAGGGGAGAGATATTTCAGTACAGCCACT 1201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2754 TGGTCCCTGGCAGAGAGAGAGCAAAATTTGAATGCCAAAATATATGTCATAGGAAT 2813
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1202 ACAATGAAAAGATAGATTTCCCAACAATATTTGCAATGTCTCAAGTTAAGCCACTG 1261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2814 TCGAT-----GATGACACTTACGACATGACATTCGCTGCGACGCTGAATCGAAT 2867
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1262 ATGATCACTGCTCTAGAAATCCAAATACGTGAAGACTGTGTGCTTGCTGATGGTCT 1321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2868 CGTCCGCTGCGCAGAGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2927
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1322 TTGCTT-----CTGGGAGTGAATGCCACATCTCTGCTGGTGGTGTTCACAAA 1369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2928 TGCAGCTGCCGAGCTGAGACGATGTAGCTCTCGGGCTAGAGGCAAGATAGGCTTCT 2987
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1370 CAGGAAAAGGCTCCGCCGACCTCTGATGCCAAATCAAGTCAAGTCCACACTTTGT 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2988 CTCCTTTATTTCGAGCGGCTGAGAGAGGCTCATGTACACTGATCCACATCCAGCCGT 3047
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1430 GCACATCCGCCACACTATATGACCATGATGATGATGATGATGATGATGATGATGAT 1489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3048 GCGATCACAACATTTACTTACAGAACACTGACCCACACATGCTGTGTCTGGAACCA 3107
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1490 TTCAGAAACCTGG-----GCAAGACACTGCCAGGCTGACCTCTGAGAGCC 1534
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3108 CTCGGACGGGCGGCCAGGCAAACTTGCACGACCGCTGCCAGGCGATTCGGAGGCG 3167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1535 CCGTGAAGCTGTGAAGAGGAGGCACTACTACGCTATGAGATGAGAGTGGGCGCTGG 1594
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3168 CCGTGTGTGTCTGAACGATGGCCGACATGACTTGGTGGGCAATCATGCTGAGGCGCTGG 3227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1595 AGTGTGGGAAGG-----CCAGGGGTCTACACCCAGATTACCAATTTCTGAATTGA 1648
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3228 GCTGTGACAGAGATGTCCGGGTGTGACACCAAGGTTCCCAACTACCTAGACTGA 3287
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1649 TCAGAACCCACCAT 1661
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3288 TTCGTGACACAT 3300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14

```

: PCT-US95-09576-1
: Sequence 1, Application PC/TUS9509576
: GENERAL INFORMATION:
:   APPLICANT: GENENTECH, INC.
:   TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
:   NUMBER OF SEQUENCES: 4
:   CORRESPONDENCE ADDRESS:
:   ADDRESS: Genentech, Inc.
:   STREET: 460 Point San Bruno Blvd
:   CITY: South San Francisco
:   STATE: California

```

```

: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/09576
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/286740
: FILING DATE: 05-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee, Wendy M.
: REGISTRATION NUMBER: 00,000
: REFERENCE/DOCKET NUMBER: 798PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7360 bases
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
PCT-US95-09576-1

Query Match          4.5%; Score 75.4; DB 5; Length 7360;
Best Local Similarity 50.8%; Pred. No. 4.4e-12;
Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

1028 AGGGCCACTTCTGTGTGGGCGCTGATCCACCCCTGGTGGTGTCACTGCTGCCACT 1087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2634 AGGGTTCCTGTGGGGGGGCATATCATAGCTCCCTGGATTCCTCTGCGCCGAC 2693
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1088 GCACCCA-----CATMAAACCGACATCTAAAGTGTGTAGGGGACCAGGACTGA 1141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2694 GCTTCCAGAGAGGTTTCCGCCCCACCTGACGCTGATCTTGGGAGAAACATACCGGG 2753
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1142 AGAAGAAGATTTTCATGACGACGCTTTAGGGGAGAGATATTTCAGTACAGCCACT 1201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2754 TGGTCCCTGGCAGAGAGAGAGCAAAATTTGAATGCGAAAATATGTCATAGGAAT 2813
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1202 ACAATGAAAAGATAGATTTCCCAACAATATTTGCAATGTCTCAAGTTAAGCCACTG 1261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2814 TCGAT-----GATGACACTTACGACATGACATTCGCTGCGACGCTGAATCGAAT 2867
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1262 ATGATCACTGCTCTAGAAATCCAAATACGTGAAGACTGTGTGCTTGCTGATGGTCT 1321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2868 CGTCCGCTGTCGCCAGAGAGAGAGGTGTGCTGACGTGTGTGCTTCCCGGGGAGCC 2927
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1322 TTGCTT-----CTGGGAGTGAATGCCACATCTCTGCTGGTGGTGTACAGAAA 1369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2928 TGCAGCTCCGACCTGGAGCGAGTGTGACCTCTCGGGTACGCAAGATGAGGCGTGT 2987
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1370 CAGGAAAAGGCTCCGCCGACCTCTGATGCCAAATCAAGTCAAGTCCACACTTTGT 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2988 CTCCTTTATTTCGAGCGGCTGAGAGAGGCTCATGTACACTGATCCACATCCAGCCGT 3047
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1430 GCACATCCGCCACACTTACTTACAGAACAGTCCAGACAACTGTGTGTGCGAGACA 3107
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3048 GCGATCACAACATTTACTTACAGAACAGTCCAGACAACTGTGTGTGCGAGAGGCG 3167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1490 TTCAGAAACCTGG-----GCAAGACACTGCCAGGCTGACTGTGAGGCG 1534
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3108 CTCGGACGGGCGGCCAGGCAAACTTGCACGACCGCTGCCAGGCGATTCGGAGGCG 3167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1535 CCGTGAAGCTGTGAAGAGAGGCACTACTACGCTATGAGATGAGAGTGGGCGCTGG 1594
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3168 CCGTGTGTGTCTGAACGATGGCCGACATGACTTGGTGGGCAATCATGCTGAGGCGCTGG 3227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 18:44:22 ; Search time 127.5 seconds

(without alignments)
8816.543 Million cell updates/sec

Title: US-09-912-559-1

Perfect score: 1683
Sequence: 1 atgtttgcacagatgtcttca.....aaagttaaagtgcttctaa 1683

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 478924 seqs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683	100.0	1683	10	US-09-912-559-1
2	1683	100.0	3008	10	US-09-880-107-1668
3	1679.8	99.8	1683	10	US-09-912-559-2
4	120.2	7.1	451	10	US-09-864-761-11164
5	117.6	7.0	2036	10	US-09-954-456-552
6	117.6	7.0	2036	10	US-09-880-107-1612
7	117	7.0	1117	10	US-09-864-761-27791
8	75.4	4.5	1689	10	US-09-969-271-6
9	75.4	4.5	2519	10	US-09-969-271-5
10	75.4	4.5	2519	10	US-09-974-298-144
11	72.8	4.3	614	10	US-09-879-792-33
12	68	4.0	1614	10	US-09-888-615-45
13	68	4.0	1748	10	US-09-879-792-11
14	66.4	3.9	1230	10	US-09-879-792-35
15	66.4	3.9	2412	9	US-09-878-295A-68
16	66.4	3.9	2412	9	US-09-878-697-68
17	66.4	3.9	2412	9	US-09-878-192A-68
18	66.4	3.9	2412	9	US-09-999-832A-68
19	66.4	3.9	2412	9	US-09-978-189-68

20	66.4	3.9	2412	9	US-10-174-590-63	Sequence 63, Appl
21	66.4	3.9	2412	9	US-10-176-758-63	Sequence 63, Appl
22	66.4	3.9	2412	9	US-10-175-737-63	Sequence 63, Appl
23	66.4	3.9	2412	9	US-10-173-706-63	Sequence 63, Appl
24	66.4	3.9	2412	9	US-10-175-738-63	Sequence 63, Appl
25	66.4	3.9	2412	9	US-10-175-752-63	Sequence 63, Appl
26	66.4	3.9	2412	9	US-10-176-482-63	Sequence 63, Appl
27	66.4	3.9	2412	9	US-10-176-757-63	Sequence 63, Appl
28	66.4	3.9	2412	9	US-10-176-913-63	Sequence 63, Appl
29	66.4	3.9	2412	9	US-10-180-552-63	Sequence 63, Appl
30	66.4	3.9	2412	9	US-10-180-557-63	Sequence 63, Appl
31	66.4	3.9	2412	9	US-10-173-700-63	Sequence 63, Appl
32	66.4	3.9	2412	9	US-10-174-572-63	Sequence 63, Appl
33	66.4	3.9	2412	9	US-10-174-579-63	Sequence 63, Appl
34	66.4	3.9	2412	9	US-10-174-582-63	Sequence 63, Appl
35	66.4	3.9	2412	9	US-10-174-588-63	Sequence 63, Appl
36	66.4	3.9	2412	9	US-10-175-739-63	Sequence 63, Appl
37	66.4	3.9	2412	9	US-10-175-740-63	Sequence 63, Appl
38	66.4	3.9	2412	9	US-10-175-743-63	Sequence 63, Appl
39	66.4	3.9	2412	9	US-10-176-488-63	Sequence 63, Appl
40	66.4	3.9	2412	9	US-10-176-492-63	Sequence 63, Appl
41	66.4	3.9	2412	9	US-10-176-747-63	Sequence 63, Appl
42	66.4	3.9	2412	9	US-10-176-750-63	Sequence 63, Appl
43	66.4	3.9	2412	9	US-10-176-985-63	Sequence 63, Appl
44	66.4	3.9	2412	9	US-10-176-987-63	Sequence 63, Appl
45	66.4	3.9	2412	9	US-10-176-991-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-912-559-1
Sequence 1, Application US/09912559
Patent No. US20020142316A1
GENERAL INFORMATION:
APPLICANT: ROEMISCH, JOERGEN
APPLICANT: STOEHR, HANS-ARNOLD
APPLICANT: FEUSSNER, ANNETTE
APPLICANT: LANG, WIEGAND
APPLICANT: WEIMER, THOMAS
APPLICANT: BECKER, MARGRET
APPLICANT: NERLICH, CLAUDIA
APPLICANT: MOTH-NAUMANN, GUDRUN
TITLE OF INVENTION: MUTANTS OF THE FACTOR VII -ACTIVATING PROTEASE AND
TITLE OF INVENTION: DETECTION METHODS USING SPECIFIC ANTIBODIES
FILE REFERENCE: 06478.1457
CURRENT APPLICATION NUMBER: US/09/912,559
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: DE 100 36 641.4
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: DE 100 50 040.4
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: DE 100 52 319.6
PRIOR FILING DATE: 2000-10-21
PRIOR APPLICATION NUMBER: DE 101 18 706.8
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ. ID NOS: 4
SOFTWARE: patentlin Ver. 2.1
SEQ ID NO 1
LENGTH: 1683
TYPE: DNA
ORGANISM: Homo sapiens
US-09-912-559-1

Query Match 100.0% Score 1683: DB 10: Length 1683:
Best Local Similarity 100.0% Pred. No. 0:
Matches 1683: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 ATGTTTGCACAGATGTCGATCTTCATGTCGCTGTTAAATGCTCTGTGGGAAGACA 60
|||||
DB 1 ATGTTTGCACAGATGTCGATCTTCATGTCGCTGTTAAATGCTCTGTGGGAAGACA 60

OY	61	GTGTMGGGGTTCCTCCCTGATATCTCTTTATTTGGAAAGCGCTGAGCCGAGACTGTGACCCCTGAC	120
Db	61	GGCTGGGGGTCTCCCTGATATCTCTTTATTTGGAAAGCGCTGAGCCGAGACTGTGACCCCTGAC	120
OY	121	CAGTATGATTACAGCTACGAGGATTATTAATCAGAGAGAAACACAGTACAGACTTAAAC	180
Db	121	CAGTATGATTACAGCTACGAGGATTATTAATCAGAGAGAAACACAGTACAGACTTAAAC	180
OY	181	CATGCTGAGAAATCCGAGCTGGTACTATACAGAGAGACCAAGCGATGCCATGCGAGGCCAAC	240
Db	181	CATGCTGAGAAATCCGAGCTGGTACTATACAGAGAGACCAAGCGATGCCATGCGAGGCCAAC	240
OY	241	CCCTGTGAACACGTTGGGGAGCTGCTTCGTCATGGAGGACCCCTTACATGACGTGCTG	300
Db	241	CCCTGTGAACACGTTGGGGAGCTGCTTCGTCATGGAGGACCCCTTACATGACGTGCTG	300
OY	301	GGTCCTTTTCTGGGAAATAAAGTGTGCAGAAAGTCAAAATACGTGGAAGGACCAACCCATGT	360
Db	301	GGTCCTTTTCTGGGAAATAAAGTGTGCAGAAAGTCAAAATACGTGGAAGGACCAACCCATGT	360
OY	361	GGCCGGGGCCAAATGTCTCTATTACCCAGAGTCTCCCTACTACCGCTGTGTGTAAACAC	420
Db	361	GGCCGGGGCCAAATGTCTCTATTACCCAGAGTCTCCCTACTACCGCTGTGTGTAAACAC	420
OY	421	CTTTACACAGGCTCCGACGTGCTCCAGTGGTTCCTGTATGGAGGCCCAACCCCTGCGAG	480
Db	421	CTTTACACAGGCTCCGACGTGCTCCAGTGGTTCCTGTATGGAGGCCCAACCCCTGCGAG	480
OY	481	AATGGGGCTACCTGCTCCCGCATAAAGCGAGATCCAACTTACCTGTGCCGTGCCGAC	540
Db	481	AATGGGGCTACCTGCTCCCGCATAAAGCGAGATCCAACTTACCTGTGCCGTGCCGAC	540
OY	541	CAGTTTAAAGGGGAAATTTGTGAATATAGTTATGATGATGCTATGTATGGCGATATGCTAC	600
Db	541	CAGTTTAAAGGGGAAATTTGTGAATATAGTTATGATGATGCTATGTATGGCGATATGCTAC	600
OY	601	TCTTACCAGGGGAAATAAATAGACAGCTACACACAGCTGCTGCTTTACTGGAACCTCC	660
Db	601	TCTTACCAGGGGAAATAAATAGACAGCTACACACAGCTGCTGCTTTACTGGAACCTCC	660
OY	661	CACCTCCTCTTGCAGAGAAATTACACATGTTTATGAGAGATGCTGAACCCATGGGATT	720
Db	661	CACCTCCTCTTGCAGAGAAATTACACATGTTTATGAGAGATGCTGAACCCATGGGATT	720
OY	721	GGGGAAACAATTTGTGAGAAACCCGAGTGGGAGGAGAAAGCCCTGGTGTATTATAA	780
Db	721	GGGGAAACAATTTGTGAGAAACCCGAGTGGGAGGAGAAAGCCCTGGTGTATTATAA	780
OY	781	GTTTACCAATGACAGAGTGAATTGGCAATACGTATGATGTCTAGCCTGCTCAGCCAGGAC	840
Db	781	GTTTACCAATGACAGAGTGAATTGGCAATACGTATGATGTCTAGCCTGCTCAGCCAGGAC	840
OY	841	GTTGGCTACCCAGAGAAAGCCCACTGAGACCAACCAAGCTTCGGGGTTTGACTCC	900
Db	841	GTTGGCTACCCAGAGAAAGCCCACTGAGACCAACCAAGCTTCGGGGTTTGACTCC	900
OY	901	TGTGAAAGACTGAGATAGCAGAGAGAGATCAAGAGAAATCTATGAGAGCTTTAAGAGC	960
Db	901	TGTGAAAGACTGAGATAGCAGAGAGAGATCAAGAGAAATCTATGAGAGCTTTAAGAGC	960
OY	961	ACGGGGGCAACACCTTGGCAGAGGCTCCCTCAGTCCGTGGCTTGACCACTCC	1020
Db	961	ACGGGGGCAACACCTTGGCAGAGGCTCCCTCAGTCCGTGGCTTGACCACTCC	1020
OY	1021	ATGCCCCAGGGCCACTTCTGTGTGGGGCGCTATCCACCCCTGCTGGGTGCTCACTGCT	1080
Db	1021	ATGCCCCAGGGCCACTTCTGTGTGGGGCGCTATCCACCCCTGCTGGGTGCTCACTGCT	1080
OY	1081	GGCCACTGCAACGCAATATAAAACCAAGCATCTTAAGGTGGTGTATGGGGACACGAGACTG	1140
Db	1081	GGCCACTGCAACGCAATATAAAACCAAGCATCTTAAGGTGGTGTATGGGGACACGAGACTG	1140
OY	1141	AGAGAAAGAAATTTCAAGACAGAGCTTTAGGGTGGAGAAAGATATTCAGATACACCCAC	1200

Dd	1141	AAGAAAGAATTTTCATGACGACAGCTTGAAGGGAGAAAGATTTCAGTAGCACCCAC	1200
Oy	1201	TACAAATGAAGAGATGAGATTTCCCACAATGATTAATTGACTTCACAGTTAAAGCCACTG	1260
Dd	1201	TACAAATGAAGAGATGAGATTTCCCACAATGATTAATTGACTTCACAGTTAAAGCCACTG	1260
Oy	1261	GATGGCTACTGTGCCTTAGAATAACCAATACGGAAGACTGTGTGCTTGCCCTCATGCGGTCC	1320
Dd	1261	GATGGCTACTGTGCCTTAGAATAACCAATACGGAAGACTGTGTGCTTGCCCTCATGCGGTCC	1320
Oy	1321	TTTTCCCTCTGGAGTAGTGGCCACATCTCTGGCTGGGGGTGTTCAGAAACAGGAAAAGG	1380
Dd	1321	TTTTCCCTCTGGAGTAGTGGCCACATCTCTGGCTGGGGGTGTTCAGAAACAGGAAAAGG	1380
Oy	1381	TCCCGCACCCTCTGTGATGGCCAAGAACGATGATTTGGCAACACTTTGTGCCAATCCCGC	1440
Dd	1381	TCCCGCACCCTCTGTGATGGCCAAGAACGATGATTTGGCAACACTTTGTGCCAATCCCGC	1440
Oy	1441	CAACTCTATGACCACATGATTAATGATGACAGTATGATCTGTGCAGSAAATCTTCGAAACT	1500
Dd	1441	CAACTCTATGACCACATGATTAATGATGACAGTATGATCTGTGCAGSAAATCTTCGAAACT	1500
Oy	1501	GGCAGAGACACTGGCCAGAGGACTCTGGAGGCCCCCTGCACCTGTGGAAAGAGCGCAC	1560
Dd	1501	GGCAGAGACACTGGCCAGAGGACTCTGCAGGCCCCCTGCACCTGTGGAAAGAGCGCAC	1560
Oy	1561	TACTACGTCTATGGGATATGAGAGCTGGAGCCCTGAGCTGTGGAAAGAGCGCGGCTTAC	1620
Dd	1561	TACTACGTCTATGGGATATGAGAGCTGGAGCCCTGAGCTGTGGAAAGAGCGCGGCTTAC	1620
Oy	1621	ACCAAGATTACCAAAATCTCGAATTTGATCAAAAGCCACATCAAAAGTGAATGAGCTTTC	1680
Dd	1621	ACCAAGATTACCAAAATCTCGAATTTGATCAAAAGCCACATCAAAAGTGAATGAGCTTTC	1680
Oy	1681	TAA 1683	
Dd	1681	TAA 1683	
RESULT 2			
US-09-880-107-1668			
Sequence 1668, Application US/09880107			
Patent No. US20020142981A1			
GENERAL INFORMATION:			
APPLICANT: Horne, Darci T.			
APPLICANT: Vockley, Joseph G.			
APPLICANT: Scherf, Uwe			
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer			
FILE REFERENCE: 44821-5028-WO			
CURRENT APPLICATION NUMBER: US/09/880, 107			
CURRENT FILING DATE: 2001-06-14			
PRIOR APPLICATION NUMBER: US 60/211, 379			
PRIOR FILING DATE: 2000-06-14			
PRIOR APPLICATION NUMBER: US 60/237, 054			
PRIOR FILING DATE: 2000-10-02			
NUMBER OF SEQ ID NOS: 3950			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO. 1668			
LENGTH: 3008			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
OTHER INFORMATION: Genbank Accession No. US20020142981A1 D49742			
US-09-880-107-1668			
Query Match 100.0%; Score 1683; DB 10; Length 3008;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps			
ATGTTTGGCCAGAGATGCTGATCTCCATGCTTTCGCTGTTAAATGAGCTCTGTGGGAAAGACA			
60			

Db 97 ATGTTTCCAGATGCTGATCTCCATGTTCTGCTTAATGCTCTGTGGAAAGACA 156
Qy 61 GCTGTGGGTCCCTGATGCTTTATGGAAGCTGGACCCAGACTGGACCCCTGGAC 120
Db 157 GCGTGTGGGTCTCCCTGATGCTTTATGGAAGCTGGACCCAGACTGGACCCCTGGAC 216
Qy 121 CAGTATGATTACAGCTACGAGGATTTATTAATCAGGAAGACACCACTGACACTTACC 180
Db 217 CAGTATGATTACAGCTACGAGGATTTATTAATCAGGAAGACACCACTGACACTTACC 276
Qy 181 CATGCTGAGATTCCTGACTGCTACTACACTGAGGACCAAGCTGATCCATGCGACCAAC 240
Db 277 CATGCTGAGATTCCTGACTGCTACTACACTGAGGACCAAGCTGATCCATGCGACCAAC 336
Qy 241 CCCGTGTAACAGCTGGAGGACTGCTGCTCATGGAGGACACTTACATGAGCTGCTG 300
Db 337 CCGTGTGAACAGCTGGAGGACTGCTGCTCATGGAGGACACTTACATGAGCTGCTG 396
Qy 301 GCTCCTTTCTCTGGGAAATAGTGTCAAGAAAGTCAAAATACGTGCAGACACCACTGT 360
Db 397 GCTCCTTTCTCTGGGAAATAGTGTCAAGAAAGTCAAAATACGTGCAGACACCACTGT 456
Qy 361 GCGCGGGGCCAATGCTCATTAACCAAGTCCCTGCTACTACCGGTGTGTGTAAACAC 420
Db 457 GCGCGGGGCCAATGCTCATTAACCAAGTCCCTGCTACTACCGGTGTGTGTAAACAC 516
Qy 421 CCTTACACAGCTGCCAGCTGCTCCCAAGTGTCTGTATGCAAGGCCAAACCTGCGCAG 480
Db 517 CCTTACACAGCTGCCAGCTGCTCCCAAGTGTCTGTATGCAAGGCCAAACCTGCGCAG 576
Qy 481 AATGGGGCTACCTGCTCCCGGATTAAGCGAGATCCCAAGTTCACCTGTGCTGCCGAC 540
Db 577 AATGGGGCTACCTGCTCCCGGATTAAGCGAGATCCCAAGTTCACCTGTGCTGCCGAC 636
Qy 541 CAGTTCAAGGGGAAATGCTGGAATAGTTCGATGACTGCTATGTGGAGGCTTAC 600
Db 637 CAGTTCAAGGGGAAATGCTGGAATAGTTCGATGACTGCTATGTGGAGGCTTAC 696
Qy 601 TCTTACCGAGGAAATGGAATAGCAAGTCAACCACTGCTGCTTACTGAGACTCC 660
Db 697 TCTTACCGAGGAAATGGAATAGCAAGTCAACCACTGCTGCTTACTGAGACTCC 756
Qy 661 CACCTCTCTTGCAGAGATTAACACATGTTATGAGAGATGCTGAACCACTGGGATT 720
Db 757 CACCTCTCTTGCAGAGATTAACACATGTTATGAGAGATGCTGAACCACTGGGATT 816
Qy 721 GGGGAACACATTTCTGCAGAACCCAGATGCGGACAAAGCCCTGTGCTTTATTA 780
Db 817 GGGGAACACATTTCTGCAGAACCCAGATGCGGACAAAGCCCTGTGCTTTATTA 876
Qy 781 GTTACCAATGCAAGGTGAATGGAATAGTGTCTCAGCTGCTCAGCCAGGAC 840
Db 877 GTTACCAATGCAAGGTGAATGGAATAGTGTCTCAGCTGCTCAGCCAGGAC 936
Qy 841 GTTCCCTACCGAAGAAAGCCCACTGAGCCATCAACCACTTCCGGGTGTGACTCC 900
Db 937 GTTCCCTACCGAAGAAAGCCCACTGAGCCATCAACCACTTCCGGGTGTGACTCC 996
Qy 901 TGTGGAAGAAGCTGAGTACAGAGAGAAAGATCAAGAGATCTATGAGGCTTTAAGAG 960
Db 997 TGTGGAAGAAGCTGAGTACAGAGAGAAAGATCAAGAGATCTATGAGGCTTTAAGAG 1056
Qy 961 ACGGCGGGCAAGACCATGGAGGCTCCCTCAGCTCCCTGCTCTACCATCTCC 1020
Db 1057 ACGGCGGGCAAGACCATGGAGGCTCCCTCAGCTCCCTGCTCTACCATCTCC 1116
Qy 1021 ATGCCCCAGGGCCACTTCTGTGTGGGGCTGTATCCACCCTGCTGGGTCTTACTGCT 1080
Db 1117 ATGCCCCAGGGCCACTTCTGTGTGGGGCTGTATCCACCCTGCTGGGTCTTACTGCT 1176
Qy 1081 GCCCACTGCACGACATTAACCAAGACATCTAAAGGTGCTAGGGGACACAGACTG 1140
Db 1177 GCCCACTGCACGACATTAACCAAGACATCTAAAGGTGCTAGGGGACACAGACTG 1236

Qy 1141 AAGAAGAGAAATTTATGAGCAGAGCTTTAGGTTGAGAGAAATTTCAAGTACAGCAC 1200
Db 1237 AAGAAGAGAAATTTATGAGCAGAGCTTTAGGTTGAGAGAAATTTCAAGTACAGCAC 1296
Qy 1201 TACAAATGAAGAGATGATTTCCCAATGATGATTTGATCTTCTCAAGTTAAAGCCAGTG 1260
Db 1297 TACAAATGAAGAGATGATTTCCCAATGATGATTTGATCTTCTCAAGTTAAAGCCAGTG 1356
Qy 1261 GATGTCACATGCTCTGATGAATCCAAATACGTGAAGACATGTGCTGCTGATGAGTCC 1320
Db 1357 GATGTCACATGCTCTGATGAATCCAAATACGTGAAGACATGTGCTGCTGATGAGTCC 1416
Qy 1321 TTTCCTCTGGAGTACGAGCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
Db 1417 TTTCCTCTGGAGTACGAGCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1476
Qy 1381 TCCCGCAGCTCCTGATGCCAAAGTCAAGCTGATTGCCACACTTGTGCAACTCCCGC 1440
Db 1477 TCCCGCAGCTCCTGATGCCAAAGTCAAGCTGATTGCCACACTTGTGCAACTCCCGC 1536
Qy 1441 CAACCTATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Db 1537 CAACCTATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1596
Qy 1501 GGGCAGACACCTGCGAGGCTGACTCTGAGAGGCCCTGACCTGTGAGAGAGACGACACC 1560
Db 1597 GGGCAGACACCTGCGAGGCTGACTCTGAGAGGCCCTGACCTGTGAGAGAGACGACACC 1656
Qy 1561 TACTACGCTATGAGTATGAGTACGCTGGAGGCTGAGTGTGGAGAGAGCCAGGGCTTAC 1620
Db 1657 TACTACGCTATGAGTATGAGTACGCTGGAGGCTGAGTGTGGAGAGAGCCAGGGCTTAC 1716
Qy 1621 ACCCAAGTTACCAATTTCTGAAATGATGATCAAGCCACATCAAAAGTGAAGAGCTTTC 1680
Db 1717 ACCCAAGTTACCAATTTCTGAAATGATGATCAAGCCACATCAAAAGTGAAGAGCTTTC 1776
Qy 1681 TAA 1683
Db 1777 TAA 1779

RESULT 3
US-09-912-559-2
Sequence 2, Application US/0912559
Patent No. US20020142316A1
GENERAL INFORMATION:
APPLICANT: ROEMISCH, JUERGEN
APPLICANT: STOEHR, HANS-ARNOLD
APPLICANT: FEUSSNER, ANNETTE
APPLICANT: LANG, WIEGAND
APPLICANT: WEIMER, THOMAS
APPLICANT: BECKER, MARGRET
APPLICANT: NEBLICH, CLAUDIA
APPLICANT: KUTH-NAUMANN, GUDRUN
TITLE OF INVENTION: MUTANTS OF THE FACTOR VII-ACTIVATING PROTEASE AND
TITLE OF INVENTION: DETECTION METHODS USING SPECIFIC ANTIBODIES
FILE REFERENCE: 06478.1457
CURRENT APPLICATION NUMBER: US/09/912.559
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: DE 100 36 641.4
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: DE 100 50 040.4
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: DE 100 52 319.6
PRIOR FILING DATE: 2000-10-21
PRIOR APPLICATION NUMBER: DE 101 18 706.8
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1683
TYPE: DNA

ORGANISM: Homo sapiens
US-09-912-559-2

Query Match 99.8%; Score 1679.8; DB 10; Length 1683;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGTTTGCAGAGATCTGATCTCATGTTCTGCTGTTAATGCTCTGGTGGGAAGACA 60
1 ATGTGGCCAGATCTGATCTCATGTTCTGCTGTTAATGCTCTGGTGGGAAGACA 60
61 GCGTGGGTTCTCCCTGATGTTTATGGAAGCCTGGACCCGAGCTGACCCCTGAC 120
61 GCGTGGGTTCTCCCTGATGTTTATGGAAGCCTGGACCCGAGCTGACCCCTGAC 120
121 CAGTATGATTACAGCTACGAGGATTTAATCAGGAAGACACACAGTACACCTTACC 180
121 CAGTATGATTACAGCTACGAGGATTTAATCAGGAAGACACACAGTACACCTTACC 180
181 CAGTATGATTACAGCTACGAGGATTTAATCAGGAAGACACACAGTACACCTTACC 240
181 CAGTATGATTACAGCTACGAGGATTTAATCAGGAAGACACACAGTACACCTTACC 240
241 CCGTGTGAACAGGCTGGGAGCTGCTCCATGAGGACACCTTACATGACCTGCTG 300
241 CCGTGTGAACAGGCTGGGAGCTGCTCCATGAGGAGACCTTACATGACCTGCTG 300
301 GCTCTTTCTCTGGGAATTAAGTGTGAGAAAGTGCAGAAATACGTGCAAGGACAC 360
301 GCTCTTTCTCTGGGAATTAAGTGTGAGAAAGTGCAGAAATACGTGCAAGGACAC 360
361 GCGCGGGCCATGCTCTCAATACCCAGATCTCTCTACTACCGCTGTGTCTGTAACAC 420
361 GCGCGGGCCATGCTCTCAATACCCAGATCTCTCTACTACCGCTGTGTCTGTAACAC 420
421 CCGTACACAGGCTCCAGCTGCTCCCAAGTGGTCTTATGAGGAGGACCAACCTG 480
421 CCGTACACAGGCTCCAGCTGCTCCCAAGTGGTCTTATGAGGAGGACCAACCTG 480
481 AATGGGCTCTCTGCTCCGCGCATTAAGGAGATCCCAAGTTCACCTGTGCTGCTG 540
481 AATGGGCTCTCTGCTCCGCGCATTAAGGAGATCCCAAGTTCACCTGTGCTGCTG 540
541 CAGTTCAAGGGGAATTTCTGTGAATAGTCTGTAGTACTGATGTTGGCGATGCTAC 600
541 CAGTTCAAGGGGAATTTCTGTGAATAGTCTGTAGTACTGATGTTGGCGATGCTAC 600
601 TCTTACGAGGAAATTAATGAGACAGTCAACAGCATGCTGCTTACTGGAACCTCC 660
601 TCTTACGAGGAAATTAATGAGACAGTCAACAGCATGCTGCTTACTGGAACCTCC 660
661 CACCTCCCTCTGAGAGGAAATTAACACATGTTTATGAGAGATCTGAAACCATGG 720
661 CACCTCCCTCTGAGAGGAAATTAACACATGTTTATGAGAGATCTGAAACCATGG 720
721 GGGGAACCAATTTCTGCAAAACCCAGATGGGAGCAAAAGCCCTGGTCTTATTA 780
721 GGGGAACCAATTTCTGCAAAACCCAGATGGGAGCAAAAGCCCTGGTCTTATTA 780
781 GTTACCAATGAAGGATGAATGGAAATCTGATGCTCTGAGCTGCTGACCCAGAC 840
781 GTTACCAATGAAGGATGAATGGAAATCTGATGCTCTGAGCTGCTGACCCAGAC 840
841 GTTGTCTCCAGAGGAAAGCCCATGAGGATCAACCAAGCTTCCGGGTTTGAAGCTC 900
841 GTTGTCTCCAGAGGAAAGCCCATGAGGATCAACCAAGCTTCCGGGTTTGAAGCTC 900
901 TGTGAAAGACTGAGATAGCAGAGAGAAAGATCAAGAGATCTATGGAGGCTTTAAG 960
901 TGTGAAAGACTGAGATAGCAGAGAGAAAGATCAAGAGATCTATGGAGGCTTTAAG 960
961 ACGGGGGCAAGCCCATGGAGGAGGCTCCCTCCAGTCTGCTGAGCATCTCC 1020
961 ACGGGGGCAAGCCCATGGAGGAGGCTCCCTCCAGTCTGCTGAGCATCTCC 1020

961 ACGGGGGCAAGCCCATGGAGGAGGCTCCCTCCAGTCTGCTGAGCATCTCC 1020
1021 ATGCCCAAGGCGCATTTCTGTGTGGGCGGTGATCTACCCCTCGTGGGCTGACAGCT 1080
1021 ATGCCCAAGGCGCATTTCTGTGTGGGCGGTGATCTACCCCTCGTGGGCTGACAGCT 1080
1081 GCCCAGTGCACCATTAATAAACACAGATCTAAAGGTGTGTAGGGACAGGACCTG 1140
1081 GCCCAGTGCACCATTAATAAACACAGATCTAAAGGTGTGTAGGGACAGGACCTG 1140
1141 AAGAAAGAGATTTTATGAGCAGAGCTTTAGGGTGGAGAAATATTAAGTACAGCCAC 1200
1141 AAGAAAGAGATTTTATGAGCAGAGCTTTAGGGTGGAGAAATATTAAGTACAGCCAC 1200
1201 TACATGAAGAGATGAGATTCGCCCAATGATATTCATTCCTCAAGTTAAAGCCACTG 1260
1201 TACATGAAGAGATGAGATTCGCCCAATGATATTCATTCCTCAAGTTAAAGCCACTG 1260
1261 GATGCTACGTGTCTGATGATTCGAAATACGTAAGAGCTGTGCTGCTGATGGTCC 1320
1261 GATGCTACGTGTCTGATGATTCGAAATACGTAAGAGCTGTGCTGCTGATGGTCC 1320
1321 TTTCCCTGTGGAGTGAATGATCCCATCTCTGCTGGGAGTACAGAAACAGAAAGG 1380
1321 TTTCCCTGTGGAGTGAATGATCCCATCTCTGCTGGGAGTACAGAAACAGAAAGG 1380
1381 TCCGCGCAGCTCTCTGATGCAAGTCAAGCTGATGCAACCTTGTGCAACCTCCGC 1440
1381 TCCGCGCAGCTCTCTGATGCAAGTCAAGCTGATGCAACCTTGTGCAACCTCCGC 1440
1381 TCCGCGCAGCTCTCTGATGCAAGTCAAGCTGATGCAACCTTGTGCAACCTCCGC 1440
1441 CAACTCTATGACACATGATGATGACATGATGATGATGATGATGATGATGATGATG 1500
1441 CAACTCTATGACACATGATGATGACATGATGATGATGATGATGATGATGATGATG 1500
1501 GCGCAAGACACTCTCCAGGCTGCTGAGAGGCGCCCTGACCTGTGAGAAAGGCGC 1560
1501 GCGCAAGACACTCTCCAGGCTGCTGAGAGGCGCCCTGACCTGTGAGAAAGGCGC 1560
1561 TACTAGCTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
1561 TACTAGCTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
1621 ACCCAAGTACCAATTTCTGAAATGATGATGATGATGATGATGATGATGATGATG 1680
1621 ACCCAAGTACCAATTTCTGAAATGATGATGATGATGATGATGATGATGATGATG 1680
1681 TAA 1683
1681 TAA 1683

RESULT 4
US-09-864-761-11164/C
Sequence 11164, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04

```

; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11164
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006097.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
US-09-864-761-11164

Query Match      7.1%; Score 120.2; DB 10; Length 451;
Best Local Similarity 97.6%; Pred. No. 4.9e-27;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 324 TCAGAAAGTGCAGAAATACGTGACAGACACCATGTGCGCGGGCCCAATGTCATTAC 383
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 451 TCCATAGTGCAGAAATACGTGACAGACACCATGTGCGCGGGCCCAATGTCATTAC 392

QY 384 CCAGAGTCCCTCCACTACCGCTGTCTGTAAACACCTTAAACAGTCCAGTGTG 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 391 CCAGAGTCCCTCCACTACCGCTGTCTGTAAACACCTTAAACAGTCCAGTGTG 332

QY 444 CCAG 448
    |||||
DB 331 CCAG 327

RESULT 5
; Sequence 552: Application US/09954456
; Patent No. US2002011507A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617

```

```

; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 552
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-552

Query Match      7.0%; Score 117.6; DB 10; Length 2036;
Best Local Similarity 49.8%; Pred. No. 8.2e-26;
Matches 607; Conservative 0; Mismatches 559; Indels 54; Gaps 10;

QY 469 AACCCCTGCAGAAATGAGGCTACGTGCTCCGGCATTAAGCGGATCCAAAGTCCAGT 528
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 745 ACCCTTCCCTGGAAGCGGGGACACCTGCTGATCGTGGCACCGGACACCATCTGTGT 804

QY 529 GCCTTCCCGACGATTCAAGGGAATTCGTGAATAGTTTGATCA---CTGCTAT 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 805 GCCTTCCCGACGATTCAAGGGAATTCGTGAATAGTTTGATCAAGGCTGTGTC 864

QY 586 GTTGGCGATGCTACTCTTCAAGGGAATTAAGTGAATAGTCAATCAGATGCTGTC 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 865 TTGGGGAACGCACTGTGATCGTGTGCGTGGCCAGCACCTCAGCTCGGCGCTCAGCTGC 924

QY 646 CTTTACTGGAACCTCCACCTCTCTTTCAGAGAGAAATTACAACATTTTATGAGAGTCT 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 925 CTGGCTTGAATCTCGATCTGTCTTACAGAGAGCTGCACTGAGTCTGTTGGGCGCGCG 984

QY 706 GAACCCATGGAATTTGGGGAACACAAATTTCTGAGAAACCCAGATCGGAGCAAAAGCCC 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 985 GCCTTGGGCTGGGCTGGGCGCCCATGCTACTGCGGAATCCGGAATGACAGAGAGCCC 1044

QY 766 TGGTCTTTATTAAGTTTCCAAATGACAAAGTGAATGGGAATACGTGTCAAGCC 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1045 TGGTCTTACGT---GTTGAGAGACAGCGGCTCTCTGAGATACCGCTGAGAGCC 1101

QY 826 TGTCTAGCCAGGACGTTGCTTACCCAGAGAAAGCCCACTGAGCATCAACCAAGCTT 885
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1102 TGTGATTCCTT-----CACCAAGTCCAACTGTCCACCGATCTCTGCGGACCTTG 1152

QY 886 CCGGGTTTGAATCTCTGAGAAAGACTGATAGACAGAGAGAAATCAAGAAATCTAT 945
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1153 CCTGAGCAAGCTCCCGGGGGGCGCCAGAGCTGGCGGAGAGAGACAAAGAGAGAGCTTC 1212

QY 946 GAGGCTTTAAGAGACAGCGGAGCAACCCATGAGAGAGCGTCCCTCAGTCTCGCTG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1213 CTGGGGCCAGTATCAATGAGGCTCTCTCTGCTGCGCGGCTGACCCCTG----- 1265

QY 1006 CCTGAGCAATCTCATGCCCCAGAGCACTTCTGTGTGGGGCGGTATCCACCCCTGC 1065
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1266 --GTTGGCGCATCTCATGCGGAGAGCTTGTGCGCGGAGGCTGTGCACACCTGC 1323

QY 1066 TGGGTGCTACTGTGCTGCGGCTGCAACCGCATTAACCAAGACATCTTA-----AGTGTG 1119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 1324 TGGGTGGTGTGGCCGCCCATCTGCTTCCACAGCCCCCCCCAGGACAGCGCTCCGCTG 1383
Oy 1120 GTGCTAGGGAGACCGAGACCTGTGAAGAAAGAAATTTTCATGAGCAGAGCTTTAGGGTGGAG 1179
Db 1384 GTGCTGGGGCCACTTCTTCAACCGCAGCAGCGAGCTGACGAGACCTTGGGACTGGAG 1443
Oy 1180 AAGATATTCAATGACAGCCACTACATGAAGAGATATGATTTCCACATGATATATGCA 1239
Db 1444 AAGTACATCCCGTACACCTGTACTGCTGTTCAACCCAGCGACAC---GACCTGCTC 1500
Oy 1240 TTGCTCAAGTTAAAGCCAGTGGATGTCATGCTCTAGATATCCAAATACGTGAAGACT 1299
Db 1501 CTGATCCGGCTGAAGAAAGAGGAGCGCTGTGCACACGCTCGCAGATTCCTGACAGCC 1560
Oy 1300 GTGTCCTGGCTGAT-----GGTCTTTCCTTCCTGGGAGTGAAGGCCATCTCTGGAC 1353
Db 1561 ATCTGCTGCGCCAGCCCGGACGACCTTCCCGCCAGGACACAGTGGCCATTTGCGGGCC 1620
Oy 1354 TGGGG-----TGTACAGAAACAGGAAAGGTCGCCGACGCTCTGTGATGCCAAAGTC 1407
Db 1621 TGGGGCCACTTGGATGAGAACGTGAGCGGCTACTACGCTCCCTGGGGAGGCCCTGGTGC 1680
Oy 1408 AAGCTGATTCGCAACACTTTGTCACACTCCGCGCACTCTATGACCATATGATGATGAC 1467
Db 1681 CCCCTGGTGGCCACCAAGATGACAGCCCTTGAAGTCTTACGCGCCGACATCAGCCCC 1740
Oy 1468 AGTATGATCTGTGCAGGAAATCTTCAAGAACTGGGACAGACCTGCCAGGGTACTCT 1527
Db 1741 AACATGCTGTGCTGCGGCTACTT---CGACTGCAAGTCGACCCCTGCCAGGGGACCTCA 1797
Oy 1528 GGAGGCCCCCGTGCAGTGTGAGAGGACGCACTACTACGCTATGAGGATATGATGCTGG 1587
Db 1798 GGGGGGCCCCCTGGCTGTGCAAGAAAGAGCCGTGGCTTACCTCTACGCGATCATCAGCGG 1857
Oy 1588 GGCTGAGTGTGGG-----AAGAGCCAGGGGTCTACACCCAAATTTCCAAATTTCTG 1641
Db 1858 GGTGACGGCTGCGGGCGGCTCCCAAGAGCGGGGGTCTTACACCGCGTGGCAACTATGTG 1917
Oy 1642 AATTGATCAAGCCACCAT 1661
Db 1918 GACTGATCAAGCAGCGGAT 1937

RESULT 6
US-09-880-107-1612
; Sequence 1612, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Home, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1612
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D14012
US-09-880-107-1612

Query Match 7.0%; Score 117.6; DB 10; Length 2036;
Best local similarity 49.8%; Pred. No. 8.2e-26;
Matches 607; Conservative 0; Mismatches 559; Indels 54; Gaps 10;

Oy 469 AACCCCTGCCAATGGGGCTACCTCTCCGGCATTAAGCGAGATCCAAATTTCACTGT 528
Db 745 AGCCCTTGCCTGAAGGGGGACCTCCACCTGATGTGGCCACCGGAGACACCGTGT 804
Oy 529 GCTTCCCGACCGATTCAGAGGGAATTTCTGAATATAGTCTTGATGA---CTGGAT 585
Db 805 GCTCCACACAGGCTTTCGCTGACCGCTCTGCACATCGAGCTGTATGAGCGCTTC 864
Oy 586 GTTGGCATGCTACTCTTACGAGGAAATGATAGACAGTCACACGATGCGG 645
Db 865 TTGGGAAAGGACGTGTACCGGTGGCGTGGCCAGCAGCTTGGAGCTGACGTC 924
Oy 646 CTTTACTGAACCTCCACCTCTCTTGCAGGAATTAACAATGTTTATGAGATGCT 705
Db 925 CTGGCCCTGAACTCCGATCTGTCTACAGAGAGTGCACGCTGACTCCGTGGCGCGG 984
Oy 706 GAACCCATGGGATTTGGGAAACAAATTTCTGCAGAAACCCAGATGGCGAAAGGCC 765
Db 985 GCGCTGTGGGCTGGGCCCCCATGCTCTACTGCGGAATCCGCAATGACGAGGCC 1044
Oy 766 TGGCTTTATTAAGTTTACCAATGACAAAGTGAATGGATACTGTATGTCAGCC 825
Db 1045 TGGTGTAGT---GTTGAAGAGACAGCGGCTCTCTGGAGTACTGCCCTGAGAGCC 1101
Oy 826 TGTCTAGCCGAGAGTGGCTTACCGAGAAAGCCCACTGAGCCATCAACCAAGCTT 885
Db 1102 TGGCAATCCCT-----CACCAGATCCAACTGTACCGGATCTCTGGAGACCTG 1152
Oy 886 CCGGGGTTTACTCTCTGTGAAGACTGAGATAGCAGAGAGAAATCAAGATATAT 945
Db 1153 CTTGAGCCAGCTCCCGGCGGCGGACGCTCGGCGAGGAGCACAAGAAAGAGAGCTTC 1212
Oy 946 GGAAGCTTTAAGAGCAGGCGGAGCAGCAGCAGCAGGCGCTCCAGTCTCGCTG 1005
Db 1213 CTGCGCCACGATATATGAGGAGCTCTCTCGCTCGCGGCTCGACCCCTG----- 1265
Oy 1006 CCGTACACCATCTCCATGCCCCCAGGCGCATCTGTGTGGTGGGCGCTGATCCACCCCTGC 1065
Db 1266 --GCTGGCGGCATCTACATGAGGAGACCTTGTGCGGAGACCTGTGTCACACCTGC 1323
Oy 1066 TGGGTGCTACGTGCTGCCCTGACGACATTAACCAACACATCTTA-----AGGTG 1119
Db 1324 TGGGTGTGTGCGCGCCGCTGCTTCTCCACAGCCGCCCAAGGAGACGCTCGTG 1383
Oy 1120 GTCTAGGGGACCGACCTGAGAAAGAAATTTTCATGACGAGCTTTAGGCTGGAG 1179
Db 1384 GTCTGTGGCCACACTTTTCCACCGCAGCAGCAGCGAGCTGTGCGATGAG 1443
Oy 1180 AAGATATTCAAGTACAGCCACTACATGAAGAGATGAGATTCACCAATGATATGCA 1239
Db 1444 AAGTACATCCCGTACACCTGTACTGCGTGTCAACCCAGGAGACAC---GACCTGCTC 1500
Oy 1240 TTGCTCAAGTTAAAGCCAGTGGATGTCATGCTCTAGATATCCAAATACGTGAAGACT 1299
Db 1501 CTGATCCGGCTGAAGAAAGAGGAGCGCTGTGCCACGCTCCCAATTTGTCAGCCCC 1560
Oy 1300 GTGTCCTGGCTGAT-----GGTCTTTCCTTCCTGGGAGTGAAGGCCATCTCTGGAC 1353
Db 1561 ATCTGCTGCGCCAGCCCGGACGACCTTCCCGCCAGGACACAGTGGCCATTTGCGGGCC 1620
Oy 1354 TGGGG-----TGTACAGAAACAGGAAAGGTCGCCGACGCTCTGTGATGCCAAAGTC 1407
Db 1621 TGGGGCCACTTGGATGAGAACGTGAGCGGCTACTACGCTCCCTGGGGAGGCCCTGGTGC 1680
Oy 1408 AAGCTGATTCGCAACACTTTTGCACACTCCGCGCACTCTATGACCATATGATGATGAC 1467
Db 1681 CCCCTGGTGGCCACCAAGATGACAGCCCTTGAAGTCTTACGCGCCGACATCAGCCCC 1740
Oy 1468 AGTATGATCTGTGCAGGAAATCTTCAAGAACTGGGACAGACACTGCGAGGGTACTCT 1527
Db 1741 AACATGCTGTGCTGCGGCTACTT---CGACTGCAAGTCTGACCGCTGCCAGGGGAGCTCA 1797

Db 1367 CTCCTTCTATTGGAGCGGCTGAGAGGCGTCATGTACAGCTGTAACCCATCCAGCCGCT 1426
 QY 1430 GCACCTCCCGCCCACTCTATGACCATGATTTGATGACATGATGATGTCGAGAAATC 1489
 Db 1427 GCACATCACAACATTACTTAAACAGAGTCAACCCGCAACATGCTGTGTGTGAGACA 1486
 QY 1490 TTCAGAAACCTGG-----GCAGACACCTGTCCAGGGTGTACTGTGAGGCC 1534
 Db 1487 CTCGGAGCGGGGCCCCCAGGCAAACTTTCACAGACGCTGCCAGGGCGATTGGGAGCC 1546
 QY 1535 CCTGACCTGTGAGAGGAGCGACCTACTACGTCTATGGAGTAGTGAGCUGGGGCTGG 1594
 Db 1547 CCTGGTGTGTGACAGATGAGCGCCCATGACTTGTGTGGCATCATCAGCTGGGGCTGG 1606
 QY 1595 AGGTGGGAGAGG-----CCAGGGGCTCTACACCCAGATTACCAATTCCTGAATTGGA 1648
 Db 1607 GCTGTGACAGAGAGATGTCCGGGTGTGTACACCAGGTTACCACTAGACTGGA 1666
 QY 1649 TCAAGCCACCAT 1661
 Db 1667 TTCGTGACACAT 1679

RESULT 9
 US-09-969-271-5
 : Sequence 5, Application US/09969271
 : Patent No. US20020098179A1
 : GENERAL INFORMATION:
 : APPLICANT: Pfizer Inc. (All designated states except GB and EP (SB));
 : APPLICANT: Pfizer Limited (GB and EP (GB) only)
 : TITLE OF INVENTION: Pharmaceutical Combinations
 : FILE REFERENCE: PCS10951AYPE
 : CURRENT APPLICATION NUMBER: US/09/969,271
 : CURRENT FILING DATE: 2001-10-01
 : PRIOR APPLICATION NUMBER: GB 0025473.0
 : PRIOR FILING DATE: 2000-10-17
 : NUMBER OF SEQ ID NOS: 7
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 5
 : LENGTH: 2519
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-09-969-271-5

Query Match 4.5%; Score 75.4; DB 10; Length 2519;
 Best Local Similarity 50.8%; Pred. No. 1,2e-12;
 Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

QY 1028 AGGGCCACTTCTGTGTGGGGCCCTGATCCACCCTGCTGTGGTGTCTACTGCTGCCACT 1087
 Db 1089 AGGGGTTCCTGTGGGGGCGCATCTCATCCTCTCTGTGGATTCTCTGCGGCCACT 1148
 QY 1088 GCACCGA-----CATAAACACGACATCTAAAGTGTCTGTGGGAGACGAGACTGA 1141
 Db 1149 GCTTCAGAGAGAGTTCCGCCACCACTGAGGGTATCTGGCGAAGCATACCGGG 1208
 QY 1142 AGAAGAAGATTTCATGACAGACCTTTAGGGTGGAAAGATTTCAATATACATACAGCACT 1201
 Db 1209 TGGTCCCTGGCGAGAGAGACAGAAATTTGAAGTGAATAATACATTGTCCATTAAGAAAT 1268
 QY 1202 ACAATGAAGAAGATGATTCCTCCACATGATATTGCACTTGTCAAGTTAAAGCCAGTGG 1261
 Db 1269 TCGAT-----GATGACACTTACGACATGACATTGGCCTCTGACACTGAATTCGATT 1322
 QY 1262 ATGTGTACTGTGCTCTAGATTCGAATACGTGAAGACTGTGTCTGTCCCTGATAGGTCCT 1321
 Db 1323 CGTCCCGCTGTCCAGAGAGAGCGCTGTCCGACATGTGTGCTTCCCGCGGGGAGCC 1382
 QY 1322 TTCCT-----CTGGGAGTAGTGCACATCTCGCTGGGAGTTTACAGAAA 1369
 Db 1383 TGTAGAGTGCAGAGTGAAGAGTGTGAGGTCTCCGGCTACCGCAACATGAGAGCTTGT 1442
 QY 1370 CAGGAAAGGTCGCCGACGCTCTGTGATGCCAAAGTCAAGCTGATTTGCCAACAATTGT 1429

Db 1443 CTCCTTCTATTGGAGCGGCTGAGAGGCGTCATGTACAGCTGTAACCCATCCAGCCGCT 1502
 QY 1430 GCACCTCCCGCCCACTCTATGACCATGATTTGATGACAGATGATGTCGAGAAATC 1489
 Db 1503 GCACATCACAACATTACTTAAACAGAGTCAACCCGCAACATGCTGTGTGTGAGACA 1562
 QY 1490 TTCAGAAACCTGG-----GCAGACACCTGTCCAGGGTGTACTGTGAGGCC 1534
 Db 1563 CTCGGAGCGGGGCCCCCAGGCAAACTTTCACAGACGCTGCCAGGGCGATTGGGAGCC 1622
 QY 1535 CCTGACCTGTGAGAGGAGCGACCTACTACGTCTATGGAGTAGTGAGCUGGGGCTGG 1594
 Db 1623 CCTGGTGTGTGACAGATGAGCGCCCATGACTTGTGTGGCATCATCAGCTGGGGCTGG 1682
 QY 1595 AGGTGGGAGAGG-----CCAGGGGCTCTACACCCAGATTACCAATTCCTGAATTGGA 1648
 Db 1683 GCTGTGACAGAGAGATGTCCGGGTGTGTACACCAGGTTACCACTAGACTGGA 1742
 QY 1649 TCAAGCCACCAT 1661
 Db 1743 TTCGTGACACAT 1755

RESULT 10
 US-09-974-298-144
 : Sequence 144, Application US/09974298
 : Patent No. US20020156263A1
 : GENERAL INFORMATION:
 : APPLICANT: Chen, Huel-Mei
 : TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
 : FILE REFERENCE: PA-0037 P
 : CURRENT APPLICATION NUMBER: US/09/974,298
 : CURRENT FILING DATE: 2001-10-04
 : PRIOR APPLICATION NUMBER: 60/238,331
 : PRIOR FILING DATE: 2000-05-10
 : NUMBER OF SEQ ID NOS: 194
 : SOFTWARE: PERL Program
 : SEQ ID NO 144
 : LENGTH: 2641
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc.feature
 : OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CBI
 : NAME/KEY: unsure
 : LOCATION: 2635
 : OTHER INFORMATION: a, l, c, y, or other
 : US-09-974-298-144

Query Match 4.5%; Score 75.4; DB 9; Length 2641;
 Best Local Similarity 50.8%; Pred. No. 1,2e-12;
 Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

QY 1028 AGGGCCACTTCTGTGTGGGGGCGCTGATCCACCCTGCTGTGGTGTCTACTGCTGCCACT 1087
 Db 1209 AGGGGTTCCTGTGGGGGCGCATCTCATCCTCTCTGTGGATTCTCTGCGGCCACT 1268
 QY 1088 GCACCGA-----CATAAACACGACATCTAAAGTGTCTGTGGGAGACGAGACTGA 1141
 Db 1269 GCTTCAGAGAGAGTTCCGCCACCACTGAGGGTATCTTGGGCGAAGCATACCGGG 1328
 QY 1142 AGAAGAAGATTTCATGACAGACCTTTAGGGTGGAAAGATTTCAATATACATACAGCACT 1201
 Db 1329 TGTTCCTGGCGAGAGAGAGCAAGAAATTTGAAGTGAATAATACATTGTCCATTAAGAAAT 1388
 QY 1202 ACAATGAAGAAGTGAATTCCTCCACATGATATTGCACTTGTCAAGTTAAAGCCAGTGG 1261
 Db 1389 TCGAT-----GATGACACTTACGACATGACATTGGCCTCTGACACTGAATTCGATT 1442
 QY 1262 ATGTGTACTGTGCTCTAGATTCGAATACGTGAAGACTGTGTCTGTCCCTGATAGGTCCT 1321
 Db 1443 CGTCCCGCTGTGCCGAGAGACGCTGTGTGCGCACTGTGTGCTTCCCGCGGGGAGCC 1502

Oy 1332 TTCCCT-----CTGGAGTGTAGGCGCACATCTCGGCTGGGGCTTTACAGAA 1369
 Db 1503 TGCACGCTGCCGGACATCGACAGSAGTGTAGCTCTCGGGCTACGGCAGACAGTATGAGGCTTTGT 1562
 Oy 1370 CAGAAAAAGGTCGCCCGCAGCTCCTGGATGCCAAAGTCAGCTGATTCGCAACACTTTGT 1429
 Db 1563 CTCCTTTATTTCGGAGGGGCTGAAGAGGCTCATGTGACATCTACCCATCCAGCGCT 1622
 Oy 1430 GCACACTCCCGCCACATCTATGACACATGATTATGATACGTTATGATCTGTGACAGGAATC 1489
 Db 1623 GCACATTCACACATTTACTTTAAAGAACAGTACCGACAAACATGCTGTGTCTGTGGACA 1682
 Oy 1490 TTTCGAACCTCG-----GCAGACACCTTGCAGGGTACTCTGGAGGCC 1534
 Db 1683 CTGGAGGCGGGGCGGCCCGCAGCAACTTTGCAGACAGCCCTGCTAGAGGCGATTGCGGAGGCC 1742
 Oy 1535 CCCTGACCTGTGACAGGACGGCACCTACTACTCTATGGGATAGTAGCTGAGGCGCTGG 1594
 Db 1743 CCTGTGTGTGTGTGACAGTATGCGCGCTGACATTTGTGTGGCATCATGCGTGGGGCCTGG 1802
 Oy 1595 AGCTGTGAGGAAGG-----CCAGGGTGTACACCCACAGTTACCAAAATTCCTGAATTTGA 1648
 Db 1803 GCCTGTGACAGGAAGATGTCCCGGTGTGTACACCAAGTTACCAACTACTGTAGACTGA 1862
 Oy 1649 TCAAAGCCACCAT 1661
 Db 1863 TTCTGTACACAT 1875

```

: RESULT 11
: US-09-879-792-33
: Sequence 33, Application US/09879792
: Patent No. US20020061850A1
: GENERAL INFORMATION:
: APPLICANT: Xiao, Yonghong
: TITLE OF INVENTION: Regulation of Human Transmembrane Serine
: TITLE OF INVENTION: Protease
: FILE REFERENCE: 02973.00035
: CURRENT APPLICATION NUMBER: US/09/879,792
: CURRENT FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/211,224
: PRIOR FILING DATE: 2000-06-13
: PRIOR APPLICATION NUMBER: US 60/283,353
: PRIOR FILING DATE: 2001-04-13
: PRIOR APPLICATION NUMBER: US 60/283,648
: PRIOR FILING DATE: 2001-04-16
: PRIOR APPLICATION NUMBER: PCT _____ (Docket No. US20020061850A1 LIO-81-WO)
: PRIOR FILING DATE: 2001-06-12
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 33
: LENGTH: 614
: TYPE: DNA
: ORGANISM: mouse
: US-09-879-792-33

```

Query Match	Similarity	Score	72.8	DB	10	Length	614
Best Local	Similarity	57.28	Pred.	No. 3.1e-12			
Matches	175	Conservative	0	Mismatches	122	Indels	9
				Gaps			
QY	1381	TCGCCGCACTCTCGATGTCGCAAAAGTCAGTCAGATTTGGCCAACTCTTGTGCACCTCCGC	1440				
Db	195	TCGCCCTTCCTCCGAGAGGTTCAAGTCACCTCATTTGACTTCAGAAAGTCGCAATGACTAC	254				
QY	1441	CAACTCTATGACACACATGATTGATGACACTATGATCTGTGCAGGAAATCTTCGAAACCT	1500				
Db	255	TTGGTCTATGACAGCTACTCTTACCTCCCAAGAAATGATGTGTGCCGGGAGTCTACAGAGA	311				
QY	1501	GGGACGACACCTGCGACGGGGGACTCTGGAGGCCCCCGACACTGTGAGAGGACGGGCAC	1560				
Db	312	GGGAGGGACTCTGCGCAGGAGACAGGAGGAGCACTCTGTCTGTGTGACGAACAATCC	371				

Qy	1561	TKATACGCTCTATATGGATTA	CTAGTACGCTGGGGCCCTG	CAGTGTG-----GAA	GAGCCAGGG	1612
Db	372	TGGTACTGGCAGGCTAC	CCAGCTGGGGACAGGCTTG	CGCCAGAAATACACCTGGT		431
Qy	1615	GCTACACCCAGTTACCA	AAATTCCTGAAATGGAT	CAAGCCACCATCAAAAGT		1674
Db	432	GGTAGACAAATGACGA	AGTACTTCCTGATTTACA	GAAAGTGGAGAGGAGCTA		491
Qy	1675	GAGCTTC	1680			
Db	492	CGATTC	497			

```

RESULT 12
US-09-888-615-45
: Sequence 45, Application US/09888615
: Patent No. US20020064856A1
: GENERAL INFORMATION:
: APPLICANT: PLOWMAN, GREGORY
: APPLICANT: WHITE, DAVID
: APPLICANT: CAENEPEEL, SEAN
: APPLICANT: CHANEDZAK, GLEN
: APPLICANT: MANNING, GERARD
: APPLICANT: SUDARSANAM, SUCHA
: TITLE OF INVENTION: NOVEL PROTEASES
: FILE REFERENCE: 038602/1214
: CURRENT APPLICATION NUMBER: US/09/888,615
: CURRENT FILING DATE: 2001-06-26
: PRIOR APPLICATION NUMBER: 60/214,047
: PRIOR FILING DATE: 2000-06-26
: NUMBER OF SEQ ID NOS: 150
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 45
: LENGTH: 1614
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-888-615-45

```

		4.0%:	Score 68:	DB 10:	Length 1614:
		Best Local Similarity	56.2%:	Pred. No. 1.8e-10:	
		Matches 172:	Conservative 0:	Mismatches 125:	Indels 9:
					Gaps 2
QY	1381	TCGCGCAGCTCCTGGATGTCGCAAAAGTCAAGCTGATTGGSCCAACCTTTGTCAACTCCCGC	1440		
Db	1300	TCGCGCTTCTCCCGGAGGTGCAGGCTCATCTTCATCCAGTTCCAGAAATGCAATATAC	1359		
QY	1441	CAACTCTATGACCCACATGATTGTGATGACATGTGATGTGTCAGAGAAATGTCAGAACT	1500		
Db	1360	TTGGTCTATGACATTTACCTTACCCCAAGATGATGTGTCTGGGGACCTTGCT---	1416		
QY	1501	GCGCAAGACACCTTCCAGGCTGACTCTGGAGCCCCCTGACTGTGTGAAGAGACGGCAC	1560		
Db	1417	GCGACAGACTCTTCCACAGGAGACACCGGGGGGCTCTTGTCTGTATGACGAACCAACGC	1476		
QY	1561	TACTACGCTATGGGATGATGACTGGGCGCTGAGATGTGGAAAG-----GCCACGG	1614		
Db	1477	TGGTACTGGCAGGCTGTACACACTGTGGGCAACAGGCTGTGGCCAGAGAACCAACTGTG	1536		
QY	1615	GTCCTACACCCCAAGTTACCAATTTCTGCAATTGGATCAAAACCACCTCTAAAGTGAAGT	1674		
Db	1537	GTGTACCAACCAAGTACAGAGATTTCTCCCTGGATTTCACAGCAAGATGGAGCGAGGTG	1596		
QY	1675	GGCCTTC	1680		
Db	1597	CGATTTC	1602		

RESULT 13
US-09-879-792-11
; Sequence 11, Application US/09879792
; Patent No. US20020061850A1
; GENERAL INFORMATION:

```
APPLICANT: Xiao, Yonghong
APPLICANT: Gedlich, Richard
TITLE OF INVENTION: Regulation of Human Transmembrane Serine
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/09/879,792
CURRENT FILING DATE: 2001-06-13
PRIORITY FILING DATE: 2000-06-13
PRIORITY FILING DATE: 2000-06-13
PRIORITY FILING DATE: 2001-04-13
PRIORITY FILING DATE: 2001-04-16
PRIORITY FILING DATE: 2001-04-16
PRIORITY FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1748
TYPE: DNA
ORGANISM: Homo sapiens
US-09-879-792-11
```

```
Query Match 4.0%: Score 68; DB 10; Length 1748;
Best Local Similarity 56.2%: Pred. No. 1.8e-10;
Matches 172; Conservative 0; Mismatches 125; Indels 9; Gaps 2;
```

```
QY 1381 TCCGCGCAGCTCTCGATGCCAAGTCAAGCTGATGCCACACTTGTGCACTCCCGC 1440
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1385 TCCGCTCTCTCCGAGAGTGCAGTCAATCTCATGACTTCAGAAATGCAATGACTAC 1444
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1441 CAACCTATGACACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1445 TTGGCTATGACAGTTACTTACCCCAAGATGATGATGATGATGATGATGATGATGAT 1501
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1501 GGGCAGACACCTGCCAGGGTGAAGTGGAGGCCCTGAGTGGAGAGGAGGAGC 1560
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1502 GGCAGAGACTCTCCAGGAGAGACAGGGGGGGCTCTGTCTGTGACAGAGAGAGC 1561
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1561 TACAGCTGATGAGTGTGAGCTGGGGGCTGAGTGGGGAGAGAGAGAGAGAGAGAG 1614
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1562 TGGTACTGTCGAGGTGTACCAAGCTGGGGGACAGAGCTGTGGGACAGAGAGAGAG 1621
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1615 GTCTAGACCCAGTATCAAAATTCCTGTAATTTGATCAAAAGCAGCAGCAGTGAAGT 1674
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1622 GTGACACCAAGTACAGAGATTTCTTCTGATTTACAGAGATGAGAGAGAGAGTGTG 1681
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1675 GGCTTC 1680
      |||||
DB 1682 CGATTTC 1687
```

```
RESULT 14
US-09-879-792-35
Sequence 35, Application US/09879792
Patent No. US20020061850A1
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
APPLICANT: Gedlich, Richard
TITLE OF INVENTION: Regulation of Human Transmembrane Serine
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/09/879,792
CURRENT FILING DATE: 2001-06-13
PRIORITY FILING DATE: 2000-06-13
PRIORITY FILING DATE: 2000-06-13
PRIORITY FILING DATE: 2001-04-13
PRIORITY FILING DATE: 2001-04-16
PRIORITY FILING DATE: 2001-04-16
PRIORITY FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 36
(Docket No. US20020061850A1 LIO-81-WO)
```

```
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 35
LENGTH: 1230
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(1230)
OTHER INFORMATION: n = A,T,C or G
US-09-879-792-35
```

```
Query Match 3.9%: Score 66.4; DB 10; Length 1230;
Best Local Similarity 56.4%: Pred. No. 4.6e-10;
Matches 168; Conservative 0; Mismatches 121; Indels 9; Gaps 2;
```

```
QY 1370 CAGGAAAGGAGTCCCGGAGCTCTGATGCCAAGTCAAGCTGATGCACTTCT 1429
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 866 CAGACAGACATCCCTCTCTCCGAGAGTGCAGTCAATCTCATGCACTTCAGAAAT 925
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1430 GCACCTCCCGCAACTCTATGACACATGATGATGATGATGATGATGATGATGATGAT 1489
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 926 GCATGACTACTGTCGTATGACAGTACTTACCCCAAGATGATGATGATGATGATGAT 985
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1490 TTCAGAAACCTGCGCAAGACCTGCGCAGGGTGCAGTCTGAGAGGCCCTGACCTGTGGA 1549
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 986 TTCTGTG--GGGACAGACTCTGCGCAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAG 1042
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1550 AGCAGCGCAGCTACTACGCTCTATGAGATAGTACGCTGGGGCTGAGTGGGAGAG-- 1607
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1043 ACAACACCGCTGTGATGCTGACAGTGTACACAGTGTACACAGCTGGGGCAGAGAGAA 1102
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1608 ----GCGAGGGGCTGCTACCCCAAGTACCAAAATTCGAAATGGAATCAAGCCAGCAT 1661
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1103 ACAACCTGATGATGACCAAGTGAAGAGTTCTTCCCTGATTTACAGCAGAGAT 1160
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 15
US-09-978-295A-68
Sequence 68, Application US/09978295A
Patent No. US20020156006A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P011
CURRENT APPLICATION NUMBER: US/09/978,295A
CURRENT FILING DATE: 2001-10-15
```



```
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697.
```

Query Match 3.9% Score 66.4; DB 9; Length 2412;

Best Local Similarity 54.6%; Pred. No 7, 1e-10; Mismatches 156; Indels 12; Gaps 3;

```
Matches 202; Conservative 0; Mismatches 156; Indels 12; Gaps 3;

QY 1310 CTGATGGGTCCTTCCCTCGGAGTGCACATCTCTGGCTGGGTTGTACAGAAA 1369
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1108 CTGAAGAGACTTCCCGATGGAAAGTGTCTGACGTGATGGGGGCCACAGAG 1167
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1370 CAGGAAAGGTCGCCGCACTCTG---GATGCCAAGTCAAGCTGATTGCCACACTT 1426
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1168 ATGGAGGTGACGCTCCCTCTCTGACACACGCGGCCGCTTGCATTTCACACAGA 1227
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1427 TGTGCAACTCCCGCAACTCTATGACCAATGATTGATGACAGTATGATCTGTGAGAA 1486
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1228 TCTGCAACACAGGACGTGTAGGTGGATCATCTCCCTCATGCTGTGCGGCT 1287
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1487 ATCTCAGAAACCTGGGCAAGACACTGCCAGGCTGACTCTGGAGGCCCTTGACCTGTG 1546
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1288 ACCT---GACGGGTGGGTGTGACGTGTCCAGGGGACACGGGGGCCCTGTGTCTC 1344
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1547 AGAGGACGGCACCCTACTACTGTATGGGATAGTGAAGTGGGCCCTGAGAGTG-----TG 1600
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1345 AAGAGAGAGGCTGTGTGAGTTAGTGGAGCCACCAGCTTTGGCATGGCTGCGCAGAGG 1404
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1601 GGAAGAGGCCAGGGGTCTACACCAGTACCAATTCCTGAATTGGATCAAGCCACCA 1660
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1405 TGAACAGCCTGGGGTGTACACCCGTCTACCTCTCTGAGCTGATCCACAGAGAGA 1464
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1661 TCAAAAGTGA 1670
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1465 TGGAGAGAGA 1474
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: March 6, 2003, 23:07:00
Job time : 139.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 12:17:47 ; Search time 4285 Seconds

(without alignments)
11430.577 Million cell updates/sec

Title: US-09-912-559-2

Sequence: 1 atgttcgcaagatgtctga.....aaagtgaagtgtctctaa 1683

Scoring table: IDENTITY_NJC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

GenBml: *
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vi:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683	100.0	1683	6	AX383955
2	1679.8	99.8	1683	6	AX383954
3	1679.8	99.8	2408	6	S83182
4	1679.8	99.8	3008	6	AX409021
5	1679.8	99.8	3008	9	HUMHGFAL
6	1678.2	99.7	2251	9	BC031412
7	1047	62.2	2119	10	BC031775
8	261.2	15.5	188349	9	AL390197
9	258	15.3	157750	2	AC006097
10	241	14.3	397	6	AX305372
11	167.4	9.9	200960	2	AC115771
12	129.8	7.7	2035	10	AF099017
13	129.8	7.7	2063	10	AF224724
14	128.2	7.6	2067	10	BC019376
15	119	7.1	68586	2	AC106236
16	114.4	6.8	2033	6	E07591
17	114.4	6.8	2033	6	I15469
18	114.4	6.8	2033	6	I69315
19	114.4	6.8	2036	6	AX33070
20	114.4	6.8	2036	6	AX408965
21	114.4	6.8	2036	9	HUMHGFAP
22	99.8	5.9	2228	4	BTUPA
23	99.8	5.9	2355	4	BOVUKPA
24	99	5.9	2375	4	SSUPAR
25	94.6	5.6	1454	10	RNUPAM
26	94.6	5.6	2366	10	RNUPPLAC
27	93.2	5.5	2299	10	MMURKR
28	93	5.5	970	6	E07615
29	93	5.5	970	6	I15462
30	93	5.5	970	6	I69308
31	87.6	5.2	2245	4	VMPPA1A
32	87.6	5.2	2252	6	A07242
33	87.6	5.2	2252	6	A07250
34	87.6	5.2	2266	4	VMPPA1S
35	87.4	5.2	2512	6	A19618
36	87.2	5.2	2445	10	RATPATISS
37	84.8	5.0	2130	4	VMPBAB
38	84.8	5.0	2137	6	A07246
39	81.2	4.8	1620	6	I08184
40	81.2	4.8	1725	4	VMPTPA
41	81.2	4.8	2236	4	VMPPA2A
42	81.2	4.8	2257	6	A07244
43	77.6	4.6	2460	4	AC364605
44	77.2	4.6	919	6	A93672
45	77.2	4.6	2393	5	CHKUPAR

ALIGNMENTS

RESULT 1
AX383955
LOCUS AX383955 1683 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 2 from Patent EPI182258.
ACCESSION AX383955
VERSION AX383955.1 GI:19577526
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Roemisch,J., Stoehr,H.A., Feussner,A., Lang,W., Weiner,T.,
Becker,M., Nerlich,C. and Muth-Naumann,G.
Mutants of the factor vii activating protease and methods for their

detection
Patent: EP 1182258-A 2 27-FEB-2002;
Aventis Behring GmbH (DE)

FEATURES
Source /organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 441 a 437 c 435 g 370 t
ORIGIN

Query Match 100.0%; Score 1683; DB 6; Length 1683;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 AVGTTTCCAGAGATGCTGATCTCCATGTTCTGTTAATGAGCTCTGGTGGAGAACA 60
D 1 AVGTTTCCAGAGATGCTGATCTCCATGTTCTGTTAATGAGCTCTGGTGGAGAACA 60
OY 61 GCGTGGGGTTCCTCCATGCTTTATGGAAGCCTGGACCCAGACTGGACCCCTGAC 120
D 61 GCGTGGGGTTCCTCCATGCTTTATGGAAGCCTGGACCCAGACTGGACCCCTGAC 120
OY 121 CAGTATGATTACAGCTACAGAGATTATATCAGAGAGAACAGCAGTACACACTTACC 180
D 121 CAGTATGATTACAGCTACAGAGATTATATCAGAGAGAACAGCAGTACACACTTACC 180
OY 181 CATGCTGAGATCCTGACCTGCTACACAGAGACCAAGCTGATCCATGCCAGCCAC 240
D 181 CATGCTGAGATCCTGACCTGCTACACAGAGACCAAGCTGATCCATGCCAGCCAC 240
OY 241 CCCTGTGAACAGCGTGGGAGCTGCTGCTCATGGAGACACCTTGACATGACCTGCTG 300
D 241 CCCTGTGAACAGCGTGGGAGCTGCTGCTCATGGAGACACCTTGACATGACCTGCTG 300
OY 301 GCTCCTTTCTGGGGAATTAAGTGTGAGAAAGTCAAAATACGTGAGAGACACCAATGT 360
D 301 GCTCCTTTCTGGGGAATTAAGTGTGAGAAAGTCAAAATACGTGAGAGACACCAATGT 360
OY 361 GCGCGGGGCGCAATGCTCATTAACCAAGATCCATCCCTACTACCGCTGTGTGTAACAC 420
D 361 GCGCGGGGCGCAATGCTCATTAACCAAGATCCATCCCTACTACCGCTGTGTGTAACAC 420
OY 421 CCTTACACAGAGTCCAGCTGCTCCAGAGTGTCTGTATGAGAGCCCAACCCCTGACAG 480
D 421 CCTTACACAGAGTCCAGCTGCTCCAGAGTGTCTGTATGAGAGCCCAACCCCTGACAG 480
OY 481 AATGGGAGTACCTGCTCCGGCATTAAGGAGATCCAAAGTTCACCTGTGCTGCCGAC 540
D 481 AATGGGAGTACCTGCTCCGGCATTAAGGAGATCCAAAGTTCACCTGTGCTGCCGAC 540
OY 541 CAGTTCAAGGGGAAATTTCTGTGAATAGTTCATGATGATGCTATGTGGGATGAGCTAC 600
D 541 CAGTTCAAGGGGAAATTTCTGTGAATAGTTCATGATGATGCTATGTGGGATGAGCTAC 600
OY 601 TCTTACCGAGGGAATTAATGAGACAGTCAACAGACATGCTGCTTACTGGAACCTCC 660
D 601 TCTTACCGAGGGAATTAATGAGACAGTCAACAGACATGCTGCTTACTGGAACCTCC 660
OY 661 CACCTCCTCTGACAGAGATTATACACATGTTATGAGAGATGCTGGAACCACTGGGATT 720
D 661 CACCTCCTCTGACAGAGATTATACACATGTTATGAGAGATGCTGGAACCACTGGGATT 720
OY 721 GGGGAACACAAATTTCTGCAAGAAACCCAGATGCGGAGGAAACCCCTGTGTTAATAAA 780
D 721 GGGGAACACAAATTTCTGCAAGAAACCCAGATGCGGAGGAAACCCCTGTGTTAATAAA 780
OY 781 GTTACCAATGACAAAGTGAATGGGATGATGATGCTGAGCTGCTGACCCAGAGAC 840
D 781 GTTACCAATGACAAAGTGAATGGGATGATGATGCTGAGCTGCTGACCCAGAGAC 840
OY 841 GTTGCTTACCCAGAGAGAAAGCCCACTGAGCCATCAACCAAGCTTCGGGGTGTGACTCC 900
D 841 GTTGCTTACCCAGAGAGAAAGCCCACTGAGCCATCAACCAAGCTTCGGGGTGTGACTCC 900
```

```
OY 901 TGTGGAAGAGCTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGGAGGCTTTAAGAC 960
D 901 TGTGGAAGAGCTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGGAGGCTTTAAGAC 960
OY 961 AGGGGGGCAAGCAGCCCATGAGGAGGCTCCCTCCAGTCTGCGGCTCTGACACATCTCC 1020
D 961 AGGGGGGCAAGCAGCCCATGAGGAGGCTCCCTCCAGTCTGCGGCTCTGACACATCTCC 1020
OY 1021 ATGCCCAAGGGCCACTTCTGTGGTGGGCGCTGATCCACCCCTGCTGAGTGGCTCAGTCT 1080
D 1021 ATGCCCAAGGGCCACTTCTGTGGTGGGCGCTGATCCACCCCTGCTGAGTGGCTCAGTCT 1080
OY 1081 GCCCATGCTACCCGACATATAAACCAACATCTAAAGTGTGCTAGAGGAGCCAGAGACCTG 1140
D 1081 GCCCATGCTACCCGACATATAAACCAACATCTAAAGTGTGCTAGAGGAGCCAGAGACCTG 1140
OY 1141 AAGAAAGAGAAATTTCAATGAGCAGAGCTTTAGGGTGCAGAAATATTCAAGTACAGCCAC 1200
D 1141 AAGAAAGAGAAATTTCAATGAGCAGAGCTTTAGGGTGCAGAAATATTCAAGTACAGCCAC 1200
OY 1201 TACAATGGAAGAGATGAGATTCCCAATGATATTGATTCCTCAAGTTAAAGCCAGT 1260
D 1201 TACAATGGAAGAGATGAGATTCCCAATGATATTGATTCCTCAAGTTAAAGCCAGT 1260
OY 1261 GATGCTCAGCTGCTCCTAGATCAAAATACGTGAAGACCTGTGCTTGCCTGATGGCTCC 1320
D 1261 GATGCTCAGCTGCTCCTAGATCAAAATACGTGAAGACCTGTGCTTGCCTGATGGCTCC 1320
OY 1321 TTTCCCTCTGGAGATGAGTCCACATCTCTGCTGGGCTGTACAGAAACAGAGAAAGG 1380
D 1321 TTTCCCTCTGGAGATGAGTCCACATCTCTGCTGGGCTGTACAGAAACAGAGAAAGG 1380
OY 1381 TCCGCGCAGCTCTGAGTGCAGAAAGTCAAGCTGATGTCAAACTTTGTGCAACTCCGC 1440
D 1381 TCCGCGCAGCTCTGAGTGCAGAAAGTCAAGCTGATGTCAAACTTTGTGCAACTCCGC 1440
OY 1441 CAACTCTATGACCATGATTTGATGACAGATGATCTGTGCAGAAATCTTCAGAAACT 1500
D 1441 CAACTCTATGACCATGATTTGATGACAGATGATCTGTGCAGAAATCTTCAGAAACT 1500
OY 1501 GGGGCAAGACCTGCGGAGGCTGACTCTGGAGGCCCCCGGACCTGTGAGAAAGAGCGCAC 1560
D 1501 GGGGCAAGACCTGCGGAGGCTGACTCTGGAGGCCCCCGGACCTGTGAGAAAGAGCGCAC 1560
OY 1561 TACTAGCTCTATGGGATGAGCTGGGCGCTGAGAGTGTGAGAAGAGCGCAGGGCTTAC 1620
D 1561 TACTAGCTCTATGGGATGAGCTGGGCGCTGAGAGTGTGAGAAGAGCGCAGGGCTTAC 1620
OY 1621 ACCCAAGTTACCAATTTCTGAAATGGATCAAGCCACATCAAAAGTGAAGTGGCTTC 1680
D 1621 ACCCAAGTTACCAATTTCTGAAATGGATCAAGCCACATCAAAAGTGAAGTGGCTTC 1680
OY 1681 TAA 1683
D 1681 TAA 1683

RESULT 2
AX383954
LOCUS 1683 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 1 from Patent EP1182258.
ACCESSION AX383954
VERSION AX383954.1 GI:19577525
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Roemisch, J., Stoehr, H.A., Feussner, A., Lang, W., Weimer, T.,
Becker, M., Nerlich, C. and Muth-Naumann, G.
Mutants of the factor vli activating protease and methods for their
```


|||||
Db 1657 TACACGCTATAGGATAGTAGTGGGCTGAGTGGAGAAAGGAGGGGCTTAC 1716
OY 1621 ACCCAAGTACCAAAATTCCTGATGATGATCAAGCCACATCAAAAGTGAAGTGGCTTC 1680
|||||
Db 1717 ACCCAAGTACCAAAATTCCTGATGATGATCAAGCCACATCAAAAGTGAAGTGGCTTC 1776
OY 1681 TAA 1683
|||||
Db 1777 TAA 1779
|||||
RESULT 5
HMMGFAL 3008 bp mRNA linear PRI 10-FEB-1999
LOCUS HMMGFAL
DEFINITION Human mRNA for HGF activator like protein, complete cds.
VERSION D49742.1 GI:736706
KEYWORDS HGF activator like protein; serin protease.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE 1 (bases 1 to 3008)
AUTHORS Kilmura, N.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3008)
AUTHORS Kilmura, N.
TITLE Submitted
DIRECT SUBMISSION Submitted (17-MAR-1995) Naomi Kilmura, Institute for Liver
Research, Kansai Medical University; Moriyouchi, Osaka 570, Japan
(tel:06-992-1001(ex.2530) Fax:06-994-6095)
FEATURES
source Location/Qualifiers
1..3008
/organism="Homo sapiens"
/db_xref="taxon:9606"
97..1779
/codon_start=1
/product="HGF activator like protein"
/db_xref="gi:1345398"
/translation="MFRMSDLYLLMLVKTACTGSLMSLLESLDPMPTDQYD
STEDNDEBNTSTLTAEPMPTTIEDADPCQNPCEHSGDCLVIGSTFTSCAP
FSGNCKQVQNTCKDNPCGRGQCLITGSPPTVCVKHPYTGSCSVPECPNQC
NGATCSRHRKSKFTACPDQFKGKCEIGSDCVGVDGYRKNRMTVNOHACLW
NSHLLQENYNNFMEDAETHGIEHNPCENPDADEKPCWCTIKVNDKWEYCDVAC
SAODVAYPEESTPESTKLPQPDSCGTIELARERIKRLIYGFESKTAKEHWOASLOS
LPLTSMQGHRCGGALHPQWLTAAHCGDTRHLKLVYGGDGLKEEFHESQSRV
EKIPYSHYNERDELPHNDIALIKIKIPVDGHGALSKYKIVNCKLPDGSFSGSEHS
GKQVETGKSGROLIDAKVKILANTIKTICNSROLIDHMDISMICGNLORQGDTCGDD
SGGLPTCEKDGITYYIGLVSWGLEDCGRFEVITVYKFLNWKATISBSGF"
polyA_site
polyA_site 3008
BASE COUNT 791 a 837 c 703 g 677 t
ORIGIN
Query Match 99.8%; Score 1679.8; DB 9; Length 3008:
Best Local Similarity 99.98; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

|||||
Db 277 CATGCTGGAATCTCTGACTGCTGATCTACTAGTAGAGCAAGCTGATGATGCCAGCCAAC 336
OY 241 CCCCTGTAAACAGGTGGGAGCTGCTGTCATGAGAGACATTTACATGACGTGCTG 300
|||||
Db 337 CCCCTGTAAACAGGTGGGAGCTGCTGTCATGAGAGACATTTACATGACGTGCTG 396
OY 301 GCTCCCTTCTCTGGGAATAGTGTGCAAGATGCAAAATAGTGCAAGCAACCATGT 360
|||||
Db 397 GCTCTCTTCTCTGGGAATAGTGTGCAAGATGCAAAATAGTGCAAGCAACCATGT 456
OY 361 GCGCGGAGCAATGTCTCATTTACCAAGAGTCTCTCTACTACGCTGTGTGTAAAC 420
|||||
Db 457 GCGCGGAGCAATGTCTCATTTACCAAGAGTCTCTCTACTACGCTGTGTGTAAAC 516
OY 421 CCTTACACAGGTCACGCTGCTCCCAAGTGTCTCTGATGACAGGCAACCCCTGCGAG 480
|||||
Db 517 CCTTACACAGGTCACGCTGCTCCCAAGTGTCTCTGATGACAGGCAACCCCTGCGAG 576
OY 481 AATGGGCTACCTGCTCCCGGCAATAGCGGAGATCCAAAGTTACCTGTGCTCCGAC 540
|||||
Db 577 AATGGGCTACCTGCTCCCGGCAATAGCGGAGATCCAAAGTTACCTGTGCTCCGAC 636
OY 541 CAGTTCAAGGGGAAATTCGTGGAATAGGTTCTGATGACTGCTATGTTGGCATGGCTAC 600
|||||
Db 637 CAGTTCAAGGGGAAATTCGTGGAATAGGTTCTGATGACTGCTATGTTGGCATGGCTAC 696
OY 601 TCTTACAGAGGAAATGATAGACAGTCAACAGATGCGTGTCTTACTGGAATCC 660
|||||
Db 697 TCTTACAGAGGAAATGATAGACAGTCAACAGATGCGTGTCTTACTGGAATCC 756
OY 661 CACCTCTCTTGGAGAGAAATTCACATGTTTATGAGAGATGCTGAACCCATGGGAT 720
|||||
Db 757 CACCTCTCTTGGAGAGAAATTCACATGTTTATGAGAGATGCTGAACCCATGGGAT 816
OY 721 GGGGAACACATTTCTGAGAAACCAATGCGGACGAAAGCCCTGTCTTATTTAA 780
|||||
Db 817 GGGGAACACATTTCTGAGAAACCAATGCGGACGAAAGCCCTGTCTTATTTAA 876
OY 781 GTTACCAATGACAGGAAATGGAATGGAATCTGATGCTCAGCTCTCAGCCAGAC 840
|||||
Db 877 GTTACCAATGACAGGAAATGGAATGGAATCTGATGCTCAGCTCTCAGCCAGAC 936
OY 841 GTTGCTACCCAGAGAAAGCCCACTGAGCCATCAACAGCTTCCGGGGTTTGACTC 900
|||||
Db 937 GTTGCTACCCAGAGAAAGCCCACTGAGCCATCAACAGCTTCCGGGGTTTGACTC 996
OY 901 TGTGGAAGACTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGAGGCTTTAAGAC 960
|||||
Db 997 TGTGGAAGACTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGAGGCTTTAAGAC 1056
OY 961 ACGGCGGGAACACCCATGGCAAGGCGCTCCAGTCCCTGCTGCTGACCATCTCC 1020
|||||
Db 1057 ACGGCGGGAACACCCATGGCAAGGCGCTCCAGTCCCTGCTGCTGACCATCTCC 1116
OY 1021 ATGCCCCAGGGCACTTCTGTGTGTGGGCTGTGATTCACCCCTGCTGCTGCTGCT 1080
|||||
Db 1117 ATGCCCCAGGGCACTTCTGTGTGTGGGCTGTGATTCACCCCTGCTGCTGCTGCT 1176
OY 1081 GCCCAGTGCACGAGATATAAACCAGACATCTAAAGTGTGTAGGGGACCGAGACTG 1140
|||||
Db 1177 GCCCAGTGCACGAGATATAAACCAGACATCTAAAGTGTGTAGGGGACCGAGACTG 1236
OY 1141 AAGAAAGAGAAATTTCTATGAGCAGAGCTTATAGGTGCAGAGATATTCAGTAGACCCAC 1200
|||||
Db 1237 AAGAAAGAGAAATTTCTATGAGCAGAGCTTATAGGTGCAGAGATATTCAGTAGACCCAC 1296
OY 1201 TACAATGAAGAGATGAGATCCCAATGATATGATGCTGCTCAAGTAAAGCCAGTG 1260
|||||
Db 1297 TACAATGAAGAGATGAGATCCCAATGATATGATGCTGCTCAAGTAAAGCCAGTG 1356
OY 1261 GATGATGCTGCTGCTGATGATTCCAATACGTGAAGACTGTGCTGCTGCTGATGAGTCC 1320
|||||


```

QY 661 CACCTCTCTTGCAGAGAAATTACAACATGTTTATGAGAGATGCTGAAAACCATGGGATT 720
    |||||
Db 741 CACCTCTCTTGCAGAGAAATTACAACATGTTTATGAGAGATGCTGAAAACCATGGGATT 800
QY 721 GGGGAACACATTTTCTGCAGAAACCAGATGGGAGGAAAAAGCCCTGGCTTATTTAA 780
    |||||
Db 801 GGGGAACACATTTTCTGCAGAAACCAGATGGGAGGAAAAAGCCCTGGCTTATTTAA 860
QY 781 GTTACCAATGACAAAGTGAATGGAATAGTGTCTAGCTGCTGAGCCAGAGAC 840
    |||||
Db 861 GTTACCAATGACAAAGTGAATGGAATAGTGTCTAGCTGCTGAGCCAGAGAC 920
QY 841 GTTGTCTACCCAGAGGAAAGCCCATGAGCCATCAACCAAGCTCCGGGGTTTGCTCC 900
    |||||
Db 921 GTTGTCTACCCAGAGGAAAGCCCATGAGCCATCAACCAAGCTCCGGGGTTTGCTCC 980
QY 901 TGTGGAAGACTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGAGGCTTTAAGAGC 960
    |||||
Db 981 TGTGGAAGACTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGAGGCTTTAAGAGC 1040
QY 961 AGGGGGGCAAGACACCCATGGAGGGGCTCCAGTCTGCTGCTGACATCTCC 1020
    |||||
Db 1041 AGGGGGGCAAGACACCCATGGAGGGGCTCCAGTCTGCTGCTGACATCTCC 1100
QY 1021 ATGCCCCAGGGCCACTTCTGTGTGGGGCGCTGATCCACCCCTGCTGGTCTCACTGCT 1080
    |||||
Db 1101 ATGCCCCAGGGCCACTTCTGTGTGGGGCGCTGATCCACCCCTGCTGGTCTCACTGCT 1160
QY 1081 GCCCACTGACGACATAAAAACACAGCATCTAAAGTGGTGGTAGGGAGCCAGAGCTG 1140
    |||||
Db 1161 GCCCACTGACGACATAAAAACACAGCATCTAAAGTGGTGGTAGGGAGCCAGAGCTG 1220
QY 1141 AAGAAAGAAATTTTATGAGAGAGAGCTTTAGGGTGCAGAAAGATTTCAAGTACACCCAC 1200
    |||||
Db 1221 AAGAAAGAAATTTTATGAGAGAGAGCTTTAGGGTGCAGAAAGATTTCAAGTACACCCAC 1280
QY 1201 TACATGAAGAGATGAGATTTCCCAATGATTTGCTCAAGTTAAGCCAGTG 1260
    |||||
Db 1281 TACATGAAGAGATGAGATTTCCCAATGATTTGCTCAAGTTAAGCCAGTG 1340
QY 1261 GATGGTCACTGGCTTATGAGATCCAAATACGGAAGACTGGTGGTGGTGGTGGTGGTGG 1320
    |||||
Db 1341 GATGGTCACTGGCTTATGAGATCCAAATACGGAAGACTGGTGGTGGTGGTGGTGGTGGTGG 1400
QY 1321 TTTTCTCTGGGAGTGAATGACACATCTTGGCTGGGGTGTACAGAAACAGGAAAGG 1380
    |||||
Db 1401 TTTTCTCTGGGAGTGAATGACACATCTTGGCTGGGGTGTACAGAAACAGGAAAGG 1460
QY 1381 TCCCGGACACTGCTGGAGTCCCAAGTCAAGCTGATGGCAACACTTGTGCACTCCCG 1440
    |||||
Db 1461 TCCCGGACACTGCTGGAGTCCCAAGTCAAGCTGATGGCAACACTTGTGCACTCCCG 1520
QY 1441 CAACCTATGACACATGATGATGACAGATGATCTGTGACAGAAATCTTCAAGAAACCT 1500
    |||||
Db 1521 CAACCTATGACACATGATGATGACAGATGATCTGTGACAGAAATCTTCAAGAAACCT 1580
QY 1501 GGGCAAGAACTGCGCAGAGGTGACTCGAGAGGCCCTGACCTGGAGAAAGAGCCGAC 1560
    |||||
Db 1581 GGGCAAGAACTGCGCAGAGGTGACTCGAGAGGCCCTGACCTGGAGAAAGAGCCGAC 1640
QY 1561 TACTACGCTATGAGATGAGTGAAGCTGGGGCCTGAGCTGTGAGAAAGAGCCGAGGGTCTAC 1620
    |||||
Db 1641 TACTACGCTATGAGATGAGTGAAGCTGGGGCCTGAGCTGTGAGAAAGAGCCGAGGGTCTAC 1700
QY 1621 ACCCAGTTACCAATTCCTGAATTTGATCAAGCCACCATCAAAAGTGAAGAGGCTTC 1680
    |||||
Db 1701 ACCCAGTTACCAATTCCTGAATTTGATCAAGCCACCATCAAAAGTGAAGAGGCTTC 1760
QY 1681 TAA 1683
    |||
Db 1761 TAA 1763

```

RESULT 7
BC031775
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

BC031775 2119 bp mRNA linear ROD 07-AUG-2002
Mus musculus, clone MGC:28705 IMAGE:4242577, mRNA, complete cds.
BC031775
BC031775.1 GI:21594554
MGC
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2119)
Strausberg, R.
Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590.
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunsarath, P.H., Garcia, A.M., Lu, X., Huiyk, S.M., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRMA Plate: 37 Row: n Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers

1. 2119
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="PVB/N"
/clone="MGC:28705 IMAGE:4242577"
/tissue_type="Kidney, normal, 5 month old male mouse."
/clone_lib="NCI_CGAP_Kid14"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
59..1612
/codon_start=1
/product="Unknown (protein for MGC:28705)"
/protein_id="AAH31775.1"
/db_xref="GI:21594555"
/translation="MLVFRVLLALIAVGSVIGLSLMSFIAPPDPPGCSNPEHGSD
CIRIDRTFSCGPAFSGSRGCRQAKNDKNDKIDITOKHPIYRACVAPYTG
DCKVLPACRPNPCNGGCSRRHRSRFTCAPOYRGKFCIEGDDCYVDGYSYR
GKVSIVNONPCILWNSHLLLOETYNMEDAETHGLAENHPCRPDGDHRCMPCVYK
NSEKIKREYCVYCTVPPDTPNPVSESLAPWVEIDPFESGCKTEVAHVKRLVGGFK
STAGNHPWVSLQTSPLTTPSMQGRFCGALIHPCWVLTAAHCTDILKRLKLVIGD
QDKRTSEHEDTFRVERILKLSOYNEDRDLPNDAIKLKPVGHCALERSRYVTV
LPSDPFPGTGHCHTISGNGVTEGSRQLDRAKVLINLPICNSROYLDHFTDDSHIC
AGNLRPGSDPCGDSGSGPLTCEKDDTYVVGIVSMGEGCKRPQVYTVTRFLMWIK
TMHEHAGL"

CDS
BASE COUNT 577 a 561 c 530 g 451 t
ORIGIN

Query Match 62.2%; Score 1047; DB 10; Length 2119;
Best Local Similarity 82.2%; Pred. No. 2,5e-305;
Matches 1203; Conservative 0; Mismatches 260; Indels 0; Gaps 0;
QY 221 CTGATCATGCGACACCCCTGTGAACAGGTGGAGACTCCTGCTCATGAGACA 280

```

Db      150  CAGATCCATGGCAGTCCAAACCCCTGTGAACACGGCGGGGACTGATCATCAGAGGGATTA 209
Oy      281  CTTTACATGAGAGTGGCTGGCTCTTCTCTGGGATTAAGTTCAGAAAGTGCATAATA 340
Db      210  CTTTACATGAGTGGAGTGGCCCTTCTGGGGAGCGGTGCGAGACTGCAGAAACA 269
Oy      341  CCGGAGAGCAACCCATGTGGCGGGGCAATGTCTCATACCAGAGTCCCTACT 400
Db      270  AGTCAAGAGCAACCATGTGTCTCATGTGATGTCTCATATACCAGAGAGCCCTACT 329
Oy      401  ACCGCTGTGTGTAAACACCCCTTACAGAGTCCAGCTGCTCCCAAGTGTCTGTAT 460
Db      330  ACCGCTGTGTGTAAACACCCCTTACAGAGTCCAGCTGCTCCCAAGTGTCTGTAT 389
Oy      461  GCAGGCAAAACCCCTGCGCAAAATGGGGCTACCTGCTCCGCGCAATAGCGAGATCCAAGT 520
Db      380  GCAGGCAAAACCCCTGCGCAAAATGGGGCTACCTGCTCCGCGCAATAGCGAGATCCAAGT 449
Oy      521  TCACCTGTGCTGTCCCGAGCAGTTCAAGGGGAAATTTCTGTAATAGTTCGTACT 580
Db      450  TTACCTGTGCTGTCCCGAGCAGTTCAAGGGGAAATTTCTGTAATAGTTCGTACT 509
Oy      561  CCTATTGTGGCAGTGTCTACTCTTACGAGGGAAATGAATAGGACATCAACAGCATG 640
Db      510  GTTATTGTGGTGTGATGCTACTCTTACGAGGCAAAAGTGAAGACATCAACAGCATG 569
Oy      641  CGTGCTTTTACTGGAACCTCCACCTCTCTGAGGAGGAAATTCACATCTTATGAGG 700
Db      570  CATGCTTTTACTGGAACCTCCACCTCTCTGAGGAGGAAATTCACATCTTATGAGG 629
Oy      701  ATGCTGAACCCATGGGATGGGGAAACAATTTCTGAGAAACCCAGATGGGAGCAA 760
Db      630  ATGAGAGAGACCCAGGAGTGGAGAGCAACTTCTGCAAAACCCAGATGGAGCA 689
Oy      761  AGCCCTGTGCTTTTAAAGTTACCAATGACAAGGAAATGGGAAATCTGTGATGCT 820
Db      690  AACCTGTGCTTTTCTGCAAGTGAACAGTGAAGAGTGAATGGAAATCTGTGATGCT 749
Oy      821  CAGCTCTCTAGCCCAAGAGCGTTGCTTACCAGAGAAACCCCACTAGCCATCAACA 880
Db      750  CAGTCTCTCAAGTCTGACCCCTTAAACCCAGTGAAGAACCTTGTGAGCCTGTGATG 809
Oy      881  AGCTTCCGGGGTTTGACTCTGTGAAAGACTGAGATGACAGAGGAAATCAAGAA 940
Db      810  AGCTGCCAGGGTTCAGTCTCTCGGAAAGAGAGGTAGTGAACAGCATCAAGGCTA 869
Oy      941  TCTATGAGAGCTTTAAGAGCAGCGGGGCAAGCAACCATGCGAGCGTCCCTCACT 1000
Db      870  TCTACGGGGGCTTTAAGAGCAGAGAGCAAGCAACCCGTGGAGGTCTCCGAGACCT 929
Oy      1001  CGCTGCTCTGACATCTTCATGCCCCAGGGCCACTTGTGGTGGGGCGCTGATCCACC 1060
Db      930  CACTGCGGTTGACACCTTCATGCCCCAAGGCCACTTGTGGGGCCGCTGATCCACC 989
Oy      1061  CCGTGGGGGTCACCTGTGCTGCCACATGCAACATTAACCAAGCATCTAAAGGTG 1120
Db      990  CTTGCTGGGGTCTGACGCTGACGCTGACGATTAACCAACCAAGCATCTAAAGGTG 1049
Oy      1121  TGTAGGGGACAGGAGCTGGAAGAAAGAAATTTATGAGAGAGCTTTAGGGTGAAGA 1180
Db      1050  TACTAGGGGATCAGAGCTGGAAGAAAGAGATTCATGACAGACCTTCAAGGTGAAGA 1109
Oy      1181  AGATTTTCAAGTACAGCAGTCAACAATGAAGAGATGAGTTCCCAACAATGATTTGCAT 1240
Db      1110  AATTAAGTGAAGTACAGTCAATGAAGAGATGAGTTCCCAACAATGATTTGCAT 1169
Oy      1241  TGTCAAGTTAAAGCAGTGTGATGCTGTCTAGATTCAAATACGTGAAGACTG 1300
Db      1170  TGTCAAGTTAAAGCAGTGTGATGCTGTCTAGATTCAAATACGTGAAGACTG 1229
Oy      1301  TGTGCTGCTGTGATGCTGTCTGTGAGTGAAGTCCACATCTCTGCTGGGGGTG 1360

```

```

Db      1230  TATGTTGCCACAGCACCCCTTTTCCCTCTGTGAACAGTGAAGCCACATCTGTGGGGTGG 1289
Oy      1361  TTACAGAAACAGAGAAAGAGGTCCCGCAGCTCTGTGATGCCAAAGTCAAGTGAATGCCA 1420
Db      1290  TTACAGAAACAGAGAAAGAGGTCCCGCAGCTCTGTGATGCCAAAGTGAATGCCA 1349
Oy      1421  ACACTTTGTGCAACTCCCGCCACTATGACACATGATTTATACAGTATGTATCTGTG 1480
Db      1350  ACCCTTTGTGCAACTCCCGCCACTATGACACATGATTTATACAGTATGTATCTGTG 1409
Oy      1481  CAGAAATTTTCAAGAACTGTGGCAGAGACACCTGCCAGGGTGTACTGTGAGAGCCCTGCA 1540
Db      1410  CGGGGAACCTTCAGAAAGCCCGGATCAGACACCTGCGAGGTGACTGGGGGGCCCTCTAA 1469
Oy      1541  CCGTGTGAGAGAGCGACACTTACGTATGAGGATGATGAGTGTGGGGCTGTGAGTGTG 1600
Db      1470  CTTGTGAGAGAGATGAACCTTACTACTCTTACGGGATTTGTAAAGTGGGGCCAGGAATGTG 1529
Oy      1601  AGAAGAGCCAGGSGTCTACACCCAACTTACCAATTCCTGAATTTGATCAAGACCA 1660
Db      1530  GGAAGAGCCAGCAGCTTACACTCAAGTCAACCAAGTCTGAATTTGATTAAGACCA 1589
Oy      1661  TCAAAAGTCAAGTGGCTTTCTAA 1683
Db      1590  TGCACAGGAGAGGCTGGCTCTGA 1612

RESULT 8
AL390197
LOCUS
DEFINITION
Human DNA sequence from clone Rp11-190f19 on chromosome 10,
complete sequence.
AL390197
VERSION
AL390197.19 GI:15387800
KEYWORDS
HTG.
SOURCE
Human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 188349)
REFERENCE
Tracey, A.
JOURNAL
Submitted (29-Aug-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Aug 31, 2001 this sequence version replaced gi:14596344.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats: all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; SW:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP. Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone configs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/ncp/chr10
Rp11-190f19 is from the library RPI1-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBac3.6
This sequence is the entire insert of clone Rp11-190f19 The true

```


Db 977 GTCTGCTTG6CCCTGGGCCCCCTCAGCCTTACTGCGGGAACCCAGACAGATGAGAGGCT 1036
QY 766 TGGTGCTTTTAAATTAACATGACAAAGTGAATGGAACTACTGATGCTCAGCC 825
Db 1037 TGGTGCTATGT---GGTGAAGAGCAACCACTGCTGTGGAGATTGCGCCCTGACACC 1093
QY 826 TGTCTACAGCCAGAGAGCTTGGCTTACCAGAGGAAGCCCACTGAGCATCAACCAAGCT 885
Db 1094 TGTGAATTCCTGGGCAAGATGTCACCTCCA-----AACCCGGAGATTCCTACACCCCTG 1147
QY 886 CCGGGGTTTGAATCTGTGTGAAGAGCTAGATPACAGAGAGAGATCAAGAAATCTA- 944
Db 1148 CCGGAGTACAGCCCGGCTGTGGCTCTACTGTCAGCAAGAGCAAGAGAGAGAGCTTC 1207
QY 945 TGAAGGCTTTAAGAGCAGCGGCGGAGACCCATGGAGAGCGCTCCCTCATCTCCCTG 1004
Db 1208 TTGAAGACCAAGCATCATCGGGGGCTCATCATCTTGGCTGGCTACACCCCTGGCTGGCT 1267
QY 1005 GCCTCTACCATCTCATGCTCCCGAGGCGCACTTGTGTGGGCGCTGATCCACCCCTG 1064
Db 1268 GCGATCTACATTTGGGAATA-----GCTTGTGGCGGAGGCTTGTCCATACCTG 1317
QY 1065 CTGGGCTGCTACTGCTGCGCACTGACACATTAATAAC------AGACATCTAAGGT 1118
Db 1318 CTGGGTAGTGTGTGAGCCCACTGCTTGGCCACAGCGCCCGGAGAGACAGATCAGCT 1377
QY 1119 GGTCTAGGGGAGCAGAGCACTGACAAAGAAATTTCAATGAGCAGCTTTAGGGTCA 1178
Db 1378 GGTCTAGGGGAGCAGCTTCTTACACCGCAGCAGGATGTACACAGCATTTGGCATTTGA 1437
QY 1179 GAAGATATTCAGTACAGCCACATACATGAAGAGATGAGATTCCTCCCATGATATTGG 1238
Db 1438 GAAGATATTCAGTACACCTGCTACTGCTGTTCACACCA---ACAACCATGACCTTGT 1434
QY 1239 ATTTCTCAAGTTAAAGCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
Db 1495 CTGATCTGCGCTAAAGAAAGAGAGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 1554
QY 1299 TGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1352
Db 1555 CATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1614
QY 1353 CTGGGG-----TGTACAGAAACAGAAAGGCTCCCGCACTCTCTGATGCCAAGT 1406
Db 1615 CTGGGGCAATGATGATGAAGATGTAGCAGCTACTCCACTGCTGAGAGCACTGGT 1674
QY 1407 CAAGCTGATTCACCACTTTGTGCACTCCCGCACTCTATGACCAATGATGATGA 1466
Db 1675 CCGCTTGTGTGCTGACCAAGTGTAGCAGCCAGAGATGTATGTGCTGATCAGTCC 1734
QY 1467 CAGTATGATGTGCTGAGAAATCTTCAAAACCTGGGCAAGACACTGCCAGGGTGA 1526
Db 1735 TTAACATCTCTGTGGCGGCTACTT---GACTGCAATTCGATGCTGCGAGGGGATC 1791
QY 1527 TGAAGGCGCCCTGACCTGTGAGAGAGAGCAGCACTACTAGTGTATGAGATGTAG 1586
Db 1792 AGGTGGGCGCTGTGTGAGAGAAATGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1851
QY 1587 GGGGCTGAGATGTGAG-----AAGAGCCAGGGGCTTACATCCCAATTTCTTCT 1640
Db 1852 GGGTGAATGCTGTGGGCTGCTCAACAGCAGAGTCTTACACCGGATGTGCTGCTGCT 1911
QY 1641 GAATTTGATCAAAAGCCACCAT 1661
Db 1912 GGACTGATCAAGCAGCCTAT 1932

RESULT 13
AF224724 2063 bp mRNA linear ROD 30-Apr-2001
AF224724 Mus musculus hepatocyte growth factor activator HGF mRNA, complete
DEFINITION cds.
ACCESSION AF224724

VERSION AF224724.1 GI:6980071
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2063)
AUTHORS van Adelsberg, J., Sehgal, S., Kukes, A., Brady, C., Barasch, J., Yang, J., and Huan, Y.
TITLE Activation of hepatocyte growth factor (HGF) by endogenous HGF activator is required for metanephric kidney morphogenesis in vitro
JOURNAL J. Biol. Chem. 276 (18), 15099-15106 (2001)
MEDLINE 21226753
PUBMED 11032833
REFERENCE 2 (bases 1 to 2063)
AUTHORS van Adelsberg, J., Sehgal, S., Kukes, A., Brady, C., Barasch, J., Yang, J., and Huan, Y.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-2000) Medicine and Anatomy and Cell Biology, Columbia University, 630 West 168th Street, New York, NY 10032, USA
FEATURES
source
1..2063
/organism="Mus musculus"
/db_xref="taxon:10090"
5..1966
/codon_start=1
/product="hepatocyte growth factor activator HGF"
/protein_id="AF34712.1"
/db_xref="GI:6980072"
/translation="MGROMAISLCPLPRPCFLLLLLVPRGAOPQAGRNHTPP
GPNVATPVPTLPIVLSGNVSTREAPAEETBEPSEPPSSPPGGVLTBSG
OCREPRRGKMLHSCSTSESAYRKWCATHTNDRMGACAEVTLTPVSGALIDP
CASPLCNGGCTSTHDGSHYRCSPLATFKDGTKECDERIREYREVEDHARRS
EGHYDQCGMGARCEEDHTACLSPLCLNGGCHLIVGTGVSCTGLFAGFCH
IVPEHCFLENGEYRGVASTAGSLSCLANSDLYLDELHDSVAVALLGLSPHAY
CRNDKDERFLPRLIIGSSSLPESHPLALYLGNSCAGSLVHICWVSAHAFVPC
GKRHKRTPLPRLIIGSSSLPESHPLALYLGNSCAGSLVHICWVSAHAFVPC
PPRDSIVLVGSHFNRRTDVTQRTGIEKYVTVYSVFNPNHDLVLRKKGERC
AVRSEFQPTCLPEAGSFPRTGKCOLAGMGHMEVNSVSNLEALVPLVAHKS
SPREYADISPNMLCAGFYDCSDACGDSGRLPVECKKNVATLVIGIISMDGGRIN
KPYVTYVATYVIMDRIRPKRYVAT"

BASE COUNT 438 a 630 c 576 g 419 t
ORIGIN
Query Match 7.7%; Score 129.8; DB 10; Length 2063;
Best Local Similarity 50.9%; Pred. No. 1.8e-27;
Matches 621; Conservative 0; Mismatches 547; Indels 53; Gaps 11;

QY 469 AACCCCTGCAAGATGGGGCTACTGCTCCCGGCAATAGCGAGATCCAGTTCACCTGT 528
Db 737 AGCCCATGTCTGAAGAGGACCTGCTGATGTTGGGACAGGAGGACAGGCTCTGC 796
QY 529 GCCTGCCCCGACCACTTCAGGGGAATTTCTGAAATPAGT---TCTGATGACTGCTAT 585
Db 797 ACCTGCCCTGTGGCTATGCTGGCGGCTCTGTAACATGTGTCACAGAGACAGCTGCT 856
QY 586 GTTGGGATGAGCTACTTTACCGGAGGAAATTAATAGSACAGTCAACAGCATGCTGC 645
Db 857 CTGGGAATGAGTACAGATACCGAGGCGGTGCCAGCACCGCTGCGCGGCTGAGCTGC 916
QY 646 CTTTACTGAGACTCCACCTCTCTTGCAGAGAAATTAACATGTTTATGAGAGATGCT 705
Db 917 CTGGCTGGAATTTGACTGCTCTTACAGAGACTGCACTGAGACTGAGGGCTGCTGCT 976
QY 706 GAACCCATGAGGATTTGGGGAACCAATTTTGTAGAAACCCAGATGGGAGAGAAAGCC 765
Db 977 GTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036
QY 766 TGTGCTTTATTAAGTTTACCAATGACAAAGTGAATGAGGATGCTGATGCTCAGCC 825
Db 1037 TGTGCTGCTATGT---GGTGAAGAGCAACCACTGCTGTGGAGATTGCGCCCTGACACC 1093
QY 826 TGTCTACAGCCAGAGAGCTTGGCTTACCAGAGGAAGCCCACTGAGCATCAACCAAGCT 885


```

Db 922 CTGGGCTGGAATTCGACCTGCTTACCCAGAGCTGCACGTTGACTCAGTGGCTGCTGCT 981
OY 706 GAAACCCATGAGGATGAGGAAACAAATTTCTGCAAGAAACCCAGATGGGAGCAAAAGCCC 765
Db 982 GTTCGCTGTGGCTGGGCGCCCTCAGCGCTACTAGTCCGGAACCGAGCAAGATGAGAGGCT 1041
OY 766 TGGTCTTTATTAAGTTACCAATGACAAAGTGAATGGAAATGCTGATGCTCAGCC 825
Db 1042 TGGTCTCTATGT---GGTGAAGAACAAAGCAGCTGTGTGGAGATTTGCGGCTGACAGCC 1098
OY 826 TGGTCAAGCCAGGAGCTTGGCTACCCAGAGAAAGCCCACTAGTACCATCAACCAAGCT 885
Db 1099 TGTGAATCCCTGGCCAGAGTCCACTCCCA-----AACCCGGAGATCCAGACAGCCCTG 1152
OY 886 CCGGGGTTGACCTCTGTTGSAAGAACTGATGATGACAGAGAGAAATCAAGAAATCTA- 944
Db 1153 CCGAGATCAAGCCCGGCTGTGGCTCTTACTGTGTGAGAGAGGACAAAGAGAGAGCTTC 1212
OY 945 TGGAGGCTTTAAGAGCAGGCGGGCAAGCAGCAGGAGGCTGCTCAGTCTGCT 1004
Db 1213 TTGAGAGCAGCAGATCATTCGGGGGCTCATCTTGTGCTGCTGACACCCCTGGCTGCT 1272
OY 1005 GCGCTGACCATCTCATTCGCCAGGGGCACTTGTGTGGGGGCTGATCCACCCCTG 1064
Db 1273 GCGATCTCAATTGGGAATA-----GCTTGTGTGCGGGAGGCTTGTCTATACCTG 1322
OY 1065 CTGGGTGCTCACTGCTGCCCATGACAGCACTAAACAC-----AGCATCTAAAGCT 1118
Db 1323 CTGGGTGCTCACTGCTGCCCATGACAGCACTAAACAC-----AGCATCTAAAGCT 1382
OY 1119 GGTGTAGGGAGCAGGAGCTGAAAGAAAGAAATTTTATGAGAGAGAGTGTAGGTGCA 1178
Db 1383 GGTGTAGGGAGCAGGAGCTTTCACCGCAGCAGGATGTGACAGAGCAATTGGCATTTGA 1442
OY 1179 GAAAGATTTCAAGTACGACCATCAATGAAGATGAGATTTCCCAACAGATATGTC 1238
Db 1443 GAAAGATTTGCTTACACCTGTACTGCTGTTCACACCCCA---ACACCATGACTTGT 1499
OY 1239 ATTGCTCAAGTTAAAGCAGTGTGATGCTGCTTATAGATCCAAATAGCTGAAGAC 1298
Db 1500 CTGTGATCCGGCTGAAAGAAAGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1559
OY 1299 TGTGTGCTTGGCTGATG-----GCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1352
Db 1560 CATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
OY 1353 CTGGGG-----TGTACAGAAACAGAGAAAGGCTCCCGCAGCTGCTGATGCTCAAGT 1406
Db 1620 CTGGGGCCACATGATGATGAAATGTGAGAGCTACTTCAACTGCTGCTGAGGAGCTGCT 1679
OY 1407 CAAGCTGATGCAACACTTGTGCAACTCCCGCAGCTGCTGATGATGCAATGATGATGA 1466
Db 1680 CCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1739
OY 1467 CAGTATGATCTGTGAGAAATCTAGAAACCTGGGCAACACCTGCGGAGGAGCTG 1526
Db 1740 TATCATGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1786
OY 1527 TGGAGGCCCCCTGACCTGTGAGAGAGAGGCACTTACTGCTTATGAGATGAGCTG 1586
Db 1797 AGTGTGGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1856
OY 1587 GGGCTGAGAGTGTAG-----AAGAGGCCAGGGGCTTACACCCCAAGTATCCAAATCTT 1640
Db 1857 GGGTATGAGCTGTGGGGGCTTCAACAGCAGAGATCTACACCCGTGTGCTCAATTAATGT 1916
OY 1641 GAATTTGATCAAAAGCCACAT 1661
Db 1917 GGACTGTGATCAAGACGATAT 1937

```

```

LOCUS AC106236 68586 bp DNA linear HTG 14-JUL-2002
DEFINITION Rattus norvegicus clone CH230-13J10, *** SEQUENCING IN PROGRESS
ACCESSION AC106236
VERSION AC106236.2 GI:21731435
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 68586)
Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,J., Blimace,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Bruley,M., Brown,E., Brown,M., Bryant,N.P.,
Bunay,C., Butch,P., Butte,C., Butte,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Dayla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Geo,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,U.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,L.E.,
Homsli,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,B.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoen,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisseg,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabadi,K., Morgan,M., Morris,S.,
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okwundu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Plimus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoshartari,N., Sisson,I.,
Sodergren,E., Somaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,D.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 68586)
Worley,K.C.
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 68586)
Worley,K.C.
Direct Submission
Submitted (14-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:18138757.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

```

```

Center project name: GKCO
Center clone name: CH230-13J10
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 29977 bases at least Q40
Consensus quality: 32280 bases at least Q30
Consensus quality: 34098 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.bgsc.bcm.tmc.edu/docs/Genbank.draft.data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 41 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1160: contig of 1160 bp in length
1161: gap of unknown length
1261: contig of 1305 bp in length
2766: gap of unknown length
2865: contig of 1171 bp in length
4036: gap of unknown length
4136: contig of 1242 bp in length
5378: gap of unknown length
5478: contig of 1539 bp in length
7017: gap of unknown length
7117: gap of unknown length
7118: contig of 1283 bp in length
8400: gap of unknown length
8500: contig of 1062 bp in length
9562: gap of unknown length
9662: contig of 1044 bp in length
10706: gap of unknown length
10806: contig of 1380 bp in length
12186: gap of unknown length
12286: contig of 1401 bp in length
13687: gap of unknown length
13787: contig of 137 bp in length
15164: gap of unknown length
15264: contig of 1055 bp in length
16319: gap of unknown length
16419: contig of 1110 bp in length
17529: gap of unknown length
17629: contig of 1318 bp in length
18947: gap of unknown length
19047: contig of 1082 bp in length
20129: gap of unknown length
20229: contig of 1012 bp in length
21241: gap of unknown length
21341: contig of 1276 bp in length
22617: gap of unknown length
22717: contig of 1466 bp in length
24183: gap of unknown length
24283: contig of 1596 bp in length
25879: gap of unknown length
25979: contig of 2162 bp in length
28141: gap of unknown length
28241: contig of 1559 bp in length
29800: gap of unknown length
31624: contig of 1724 bp in length
31724: gap of unknown length
32926: contig of 1202 bp in length
33026: gap of unknown length
34376: contig of 1350 bp in length
34476: gap of unknown length
36187: contig of 1711 bp in length
36287: gap of unknown length
37398: contig of 1111 bp in length
37498: gap of unknown length
39415: contig of 1917 bp in length

```

```

39416 39515: gap of unknown length
39516 41278: contig of 1763 bp in length
41279 41378: gap of unknown length
41379 43355: contig of 1977 bp in length
43356 43455: gap of unknown length
43456 45020: contig of 1565 bp in length
45021 45120: gap of unknown length
45121 47073: contig of 1953 bp in length
47074 47172: gap of unknown length
47173 48525: contig of 1352 bp in length
48526 48625: gap of unknown length
48626 50339: contig of 1714 bp in length
50340 50439: gap of unknown length
50440 52494: contig of 2055 bp in length
52495 52595: gap of unknown length
52596 54163: contig of 1569 bp in length
54164 54263: gap of unknown length
54264 56386: contig of 2123 bp in length
56387 56486: gap of unknown length
56487 58244: contig of 1758 bp in length
58245 58344: gap of unknown length
58345 60191: contig of 1847 bp in length
60192 60291: gap of unknown length
60292 62620: contig of 2229 bp in length
62621 65096: gap of unknown length
65097 65196: contig of 2476 bp in length
65197 68586: contig of 3390 bp in length.

```

```

FEATURES
    source          1..68586
                    /organism="Rattus norvegicus"
                    /db_xref="taxon:10116"
                    /clone="CH230-13J10"

```

```

BASE COUNT      15651 a 16674 c 16111 g 16135 t 4015 others
ORIGIN

```

```

Query Match          7.1%; Score 119; DB 2; Length 68586;
Best Local Similarity 82.0%; Pred. No. 5.3e-24;
Matches 137; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```

```

QY 1517 AGGCTGACCTGTGGAGGCCCCCGACCTGTGAGGAGCGGACGACTACTACCTCTATGCGA 1576
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 53672 AGGCTGACCTGTGGAGGCCCCCGACCTGTGAGGAGCGGACTACTACCTCTATGCGA 53613
QY 1577 TAGTGAAGCTGGGCTGTGAGTGTGAGAGAGCGGCTTACACCCCAAGTTACCAAT 1636
    ||||||| ||||||| || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 53612 TTCTAAGCTGGGGCCAGGAGATGTGGAGAGAGCCTGAGACTACACTCAAGTCAACAA 53553
QY 1637 TCCGTAATGTGATCAAGCCACCATCAAAAGTGAAGAGTGCCTCTAA 1683
    ||||||| ||||||| || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 53552 TCCTGATTTGATMAAACCAACCATGCACAAAGAGGCTGCTCTGA 53506

```

```

Search completed: March 6, 2003, 18:44:14
Job time : 4882 secs

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 15:20:13 ; Search time 2178.5 Seconds

(without alignments)
12511.822 Million cell updates/sec

Title: US-09-912-559-2

Perfect score: 1683

Sequence: 1 atgttgcacgagatgtctga.....aaagtgaaagtgctctctaa 1683

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154056 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_estro:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_frod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	698.2	41.5	891	13	BT761782 603046775
2	451.6	26.6	592	10	BE032018 130809 MA
3	447.2	26.6	802	13	BC972579 602841165
4	436.2	25.9	674	9	AA217892 mw54b08.r
5	435.4	25.9	751	13	BT148082 602912432
6	432	25.7	570	13	BM508620 1137h03.y

7	416.4	24.7	929	12	BT788188 602113411
8	396.4	23.6	791	13	BT332440 602980841
9	394.2	23.4	546	9	AA237499 mx10b10.r
10	392.6	23.3	608	10	AV601564 AV601564
11	390.4	23.2	902	12	BT780971 602105493
12	386	22.9	682	13	BT220028 602934913
13	384.6	22.9	929	12	BT785781 602112402
14	372.4	22.1	969	12	BT384535 602046804
15	370.6	22.0	654	10	AA475402 un65h07.y
16	370.4	22.0	587	13	BC972681 602839029
17	365.6	21.7	659	10	AA610902 un13g10.y
18	363.6	21.6	594	13	BM503097 602839029
19	347.2	20.6	715	12	BT789705 602103735
20	335.8	20.0	656	10	BM569555 BM569555
21	330.8	19.7	934	13	BT765113 603051314
22	327.8	19.5	345	14	T68666 y643e09.r1
23	317.4	18.9	486	9	AA286125 yd08g12.r
24	311.8	18.5	600	12	BT788541 602114406
25	302.4	18.0	335	14	R89458 yq01a10.r1
26	293.2	17.4	304	14	T77362 yd72g02.r1
27	279.8	16.6	521	13	BT221626 602936942
28	275.8	16.4	912	13	BC972626 602836922
29	274	16.3	366	14	T93666 y606a11.r1
30	264.4	15.7	434	14	T84369 yd37g04.r1
31	258.6	15.4	781	12	BF533788 602075308
32	256.2	15.2	458	14	R10295 yf36a12.r1
33	243.4	14.6	719	13	BT102281 60288594
34	242.8	14.4	593	9	AT035669 nd60b09.x
35	241.8	14.4	513	13	BT247736 602959379
36	241	14.3	384	14	W66981 me29h02.r1
37	241	14.3	397	9	AA002504 mg45a03.r
38	239.8	14.2	353	14	R11014 yf38g10.r1
39	235.6	14.0	489	9	AT787532 u119e04.y
40	235.4	14.0	393	14	T50880 yb88d07.r1
41	213.4	12.7	530	10	BE235171 142539 MA
42	204.8	12.2	361	9	AA108230 mp38a11.r
43	197.8	11.8	320	14	T69334 yd12111.r1
44	151.8	9.0	642	13	BT067550 BT067550
45	150.4	8.9	810	9	AT255992 nt176c01.x

ALIGNMENTS

RESULT 1
BT761782 891 bp mRNA linear EST 25-SEP-2001
LOCUS 603046775F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:518706 5',
DEFINITION mRNA sequence.
ACCESSION BT761782
VERSION BT761782
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 891)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1467 row: e column: 19
High quality sequence stop: 839.
Location/Qualifiers
1..891

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5187066"
/clone_11b="NIH_MGC_116"
/lab_host="DH10B"
/Note="Organ: pooled colon, kidney, stomach; Vector:
pcmw-sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-qr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH-MGC Library."

BASE COUNT      243 a      234 c      211 g      203 t
ORIGIN

Query Match      41.5% Score 698.2; DB 13; Length 891;
Best Local Similarity 97.2% Pred. No. 7.6e-188;
Matches 7/5; Conservative 0; Mismatches 13; Indels 9; Gaps 6;

OY 1 ATGTTTCAGAGATGCTGCATCTCCATGTTCTGCTTAATGAGCTCTGGTGGAAAGACA 60
    |||||||
DB 81 ATGTTTCAGAGATGCTGCATCTCCATGTTCTGCTTAATGAGCTCTGGTGGAAAGACA 140

OY 61 GCGGTGGGTTCCCTGATGCTTATGGAAGCCTGGACCAAGCCGCTGAC 120
    |||||||
DB 141 GCGGTGGGTTCCCTGATGCTTATGGAAGCCTGGACCAAGCCGCTGAC 200

OY 121 CAGTATGATTACAGTACGAGGATTATATCAGAGAAACACAGTAGCACACTTACC 180
    |||||||
DB 201 CAGTATGATTACAGTACGAGGATTATATCAGAGAAACACAGTAGCACACTTACC 260

OY 181 CATGCTGAGATCCTGATGCTACTACTACAGAGACCAAGCTATGACCAAGCCCAAC 240
    |||||||
DB 261 CAGCTGAGATCCTGATGCTACTACTACAGAGACCAAGCTATGACCAAGCCCAAC 320

OY 241 CCTGTGAAACGCGTGGAGAGCTGCTCTCCATGGAGACACTTCAATGACAGCTGCTG 300
    |||||||
DB 321 CCTGTGAAACGCGTGGAGAGCTGCTCTCCATGGAGACACTTCAATGACAGCTGCTG 380

OY 301 GCTCCTTCTCTGGGAATAAGTGTCAAGAAATGCAAAATACGTCAAGAGCAACCACTGT 360
    |||||||
DB 381 GCTCCTTCTCTGGGAATAAGTGTCAAGAAATGCAAAATACGTCAAGAGCAACCACTGT 440

OY 361 GCGCGGGGCAATGTCTCATTCACCAAGTCTCTCCCTACTACAGCGCTGTGTTAAAC 420
    |||||||
DB 441 GCGCGGGGCAATGTCTCATTCACCAAGTCTCTCCCTACTACAGCGCTGTGTTAAAC 500

OY 421 CCTTACAGAGTCCAGAGTCTCCCAAGTGTCTCTGATGACAGGCAACCCCTGCCAG 480
    |||||||
DB 501 CCTTACAGAGTCCAGAGTCTCCCAAGTGTCTCTGATGACAGGCAACCCCTGCCAG 560

OY 481 AATGGGCTACCTCTCTCCGGCATTAAGCGAGATCCAAAGTTCATCTGTGCTGCCAG 540
    |||||||
DB 561 AATGGGCTACCTCTCTCCGGCATTAAGCGAGATCCAAAGTTCATCTGTGCTGCCAG 620

OY 541 CAGTTCAAGGGAATTCGTGAATAAGTTCTGATGACTGCTATGTTGGCGATGGCTAC 600
    |||||||
DB 621 CAGTTCAAGGGAATTCGTGAATAAGTTCTGATGACTGCTATGTTGGCGATGGCTAC 680

OY 601 TCTTTCAGAGGAAATGATAGACAGTTC - AACGAGATGCGTCTTACTAGCAATC 659
    |||||||
DB 681 TCTTTCAGAGGAAATGATAGACAGTTC - AACGAGATGCGTCTTACTAGCAATC 740

OY 660 CC - ACCCTCTTTCAGAGAAATTAACACATGTTATGAGAG - -ATGCGAAACCCATG 715
    |||||||
DB 741 CCAGAGCTCTCTGACAGAGATTAACACATGTTATGAGAGAGATGCTTAAACCCATG 800

OY 716 GGATTTGGGAA - CACAATTTCTGCAG - AAACCCAGATGGGAGG - AAAAGCCCTGTGTC 771
    |||||||

```

```

DB 801 GGATTTGGGAAACCAATTTCTGCAGAAACCAAGATGCCGACGAAAAAGCCCTGTGTC 860
OY 772 TTTATTAAGTACCAA 788
    |||||||
DB 861 TTTATTAAGTACCA 877

RESULT 2
BE032018
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 592)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
CONTACT: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68935-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAAT
BACKWARD: GTTTCCTCAGTACGACG
Plate: 69 row: K column: 12
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
source
1..592
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_11b="MARC 1P1G"
/rclone_type="pooled"
/lab_host="DH10B"
/Note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

BASE COUNT      144 a      152 c      174 g      122 t
ORIGIN

Query Match      26.8% Score 451.6; DB 10; Length 592;
Best Local Similarity 86.0% Pred. No. 1.5e-117;
Matches 518; Conservative 0; Mismatches 69; Indels 15; Gaps 1;

OY 522 CACCTGTGCTGTGCCGACAGTTCAAGGGGAATTCGTGAAATAGTTCGTGATGACTG 581
    |||||||
DB 6 CCGGTGGGCTGTGCTGTGACCAAGTTTAAGGGGAGATTCGTGAAATAGTTCGTGATGACTG 65

OY 582 CTATGTGGCATGCGTACTCTTACCGAGGAAATGATAGACAGTCAACGACATGC 641
    |||||||
DB 66 CTATGTGGCATGCGTACTCTTACCGAGGAAATGATAGTCAACGACATGC 125

OY 642 GTGCTTTACTGGAATCCCAACCTCTCTTGCAGAGGAATTCAACATGTTTATGAGGA 701
    |||||||
DB 126 GTGCTTTACTGGAATCCCAACCTCTCTTGCAGAGGAATTCAACATGTTTATGAGGA 185

OY 702 TGTGAAACCATGGGATTTGGGAGACCAATTTCTGACAGAACCCAGATGGGAGGAAA 761
    |||||||
DB 186 TGTGAGAGCCCATGGGATTTGGGAGACCAATTTCTGACAGAACCCAGATGGGAGGAAA 245

OY 762 GCGCTGTGCTTATTAAGTATGACAAAGTGAATGGGAATGGAATGATGATGCTCTC 821
    |||||||

```



```

||||| ||| ||||||| ||| ||| ||||||| ||| ||| |||
Db 246 GCGCTGGTGTTCATTAAGTACAGTGCACAGTGAATGAGACTACTGTGACGTC 305
OY 822 AGCCTGCTCAGCCAGGACGTTGCTTACCCAGAGAAAGCCCATGAGCCATCACCA 881
Db 306 TCCCTGCTCAGC-----CCAGAGGAGAACCCCATGAACCTGTGACCC 350
OY 882 GCTTCGGGGTTTGACTCCTGTGCAAGACTGAGATACGAGAGAGAACATCAGAGAA 941
Db 351 GCTTCCCGGGTTTGACTCCTGTGGAGAGACGACAGCAGCAAGAGAGTTTCAGAGAT 410
OY 942 CTATGAGAGCTTAAAGACAGCGGCGGAGACACCATGAGAGAGGTCCTCCAGTCC 1001
Db 411 CTACGAGAGCTTAAAGACAGCGGCGGAGACACCGTGGACAGGCTCTACAGACTC 470
OY 1002 GCTGCTCTGACGACATCTCCATGCCCCAGGAGGCACTTCTGTGGGGCGTATCCACC 1061
Db 471 CTTCGATCTGACGCTCCATGCCCCAGGAGGCACTTCTGTGGGGCGGCTGATCCACC 530
OY 1062 CTGCTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1121
Db 531 CTGCTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
OY 1122 GC 1123
Db 591 GC 592

RESULT 3
BG972579 802 bp mRNA linear EST 12-JUN-2001
LOCUS 602841165F1 NCI_CGAP_kid14 Mus musculus cDNA clone IMAGE:4975384
DEFINITION 5', mRNA sequence.
ACCESSION BG972579
VERSION BG972579.1 GI:14360216
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 802)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apds-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
DNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM0967 row: a column: 17
FEATURES
location/Qualifiers
Source 1..802
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4975384"
/clone_lib="NCI_CGAP_kid14"
/lab_host="DH10B (TI phase-resistant)"
/notes="Organ: kidney; Vector: pCMV-SPOB6; Site:1; NotI;
Site:2; SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP library."
BASE COUNT 211 a 230 c 203 g 158 t
ORIGIN
Query Match 26.6%; Score 447.2; DB 13; Length 802;
Best Local Similarity 77.7%; Pred. No. 3,2e-116;
Matches 580; Conservative 0; Mismatches 158; Indels 8; Caps 3;

```

```

OY 1 ATGTTCCAGATATGTCATGCTTCATGCTGTTAATGCGCTGTGGGAGAGCA 60
Db 50 ATATTTCTAGGATGTGGTGTTCGGTCTCTCTATGCTGTGATGGAGAGTCA 109
OY 61 GCGTGTGGTTCCTCCATGATGCTTTATTGAAAGCCTGAGCCAGACTGGACCCCTGAC 120
Db 110 CTCAATTGGCTCTCACTGATGATGCTTCATATGCCCCAGACCCAGATGGACCCCTG 169
OY 121 CAGTATGATTCAGCTACGAGAGATTTATATCAGAGAGAGACACAGCAGTACCTTACC 180
Db 170 GACATTTACTACAGCTATGACAGTCCAGCCAGAGAGAGACCCAGTGTACGAGAGCC 229
OY 181 CATCTAGATATCTGACTGTACTACAGTACAGCAAGCTGATCCATGCGACCCAC 240
Db 230 ACCCTGTAGAACCCCACTGGTATCT-----ATGAAAGAGATGATCCATGCGATCCAC 283
OY 241 CCGTGTAGACAGGTTGGGAGTGCCTCTGATGAGGAGCAGCCCTTCATGCACTGCTGCTG 300
Db 284 CCGTGTAGACAGCGCGGAGTGTATCATCAGAGGAGATACCTTCAGTTGGAGCTGCCA 343
OY 301 GCTGCTTTCTGAGATTAAGTGTCAAGAAAGTGCAGAAATAGCTGACAGACACCATGT 360
Db 344 GCGCCCTTCTCGGAGAGCGGTGCCAGACTGACAAACAAATGTGACAGACACCATGT 403
OY 361 GCGCGGGCCATGTCTCATTTACCCAGATCTCTCTACTACCGCTGTGTCTGTAACAC 420
Db 404 GTCCATGATGATTTGCCCTCATTTACCCAGAGACCCCTACTACCGCTGTGCTGCAATAC 463
OY 421 CCTTACACAGGTCCTCAGCTGCTCCCAAGTGTTCTGTTATGAGAGCCCAACCCCTGCCAG 480
Db 464 CCTTACACAGGACACAGACTGCTCCCAAGTGTTCTGCGCATGAGGCCCAACCCCTGCCAG 523
OY 481 AATGGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 524 AATGGGAGTGTGTTCCGACACAGAGGAGATCCAGGTTTACTGTGCTGCTGCTGCTG 583
OY 541 CAGTCAAGGGGAATTTCTGTGAA-ATAGGTTCTGATGACTCTATGTTGGCGATGGCA 599
Db 584 CAGATTAAGGGGAATTTCTGTGAACTAGGTCGAGAGCTGTATGCTGCTGATGGCTCA 643
OY 600 CTCTTACGAGGGAATTAATGATGAGGAGTGCACGACGATGGTGCCTTTACTGGAAC 659
Db 644 CTCTTACGAGGGAATTAATGATGAGGAGTGCACGACGATGGTGCCTTTACTGGAAC 702
OY 660 CCACCTCTCTGCGAGAGAAATTTACACATGTTTATGAGAGTGTGAACCCATGGGAT 719
Db 703 CCACCTCTCTGCGAGAGAAATTTATTAACATGTTTATGAGAGTGTGAACCCATGGGAT 762
OY 720 TGGGGAACACAATTTCTGCAAAACC 745
Db 763 GCGAGAGCACCACTTCTGCAAAACC 788

RESULT 4
AA217892 674 bp mRNA linear EST 06-FEB-1997
LOCUS AA217892
DEFINITION IM54D08.r1 Soares mouse 3MME12.5 Mus musculus cDNA clone
IMAGE:558839 5' similar to TR:G1345398 G1345398 HGF ACTIVATOR LIKE
PROTEIN. ; mRNA sequence.
ACCESSION AA217892
VERSION AA217892.1 GI:1826875
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 674)
AUTHORS Marfa,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Treising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

```

TITLE	The WashU-HM1 Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Marra M/Mouse EST Project

BASE COUNT	182 a	172 c	185 g	135 t
ORIGIN				

Db 421 TACTAGGGGATCTACAGACCTTGAAGAACAGATCCCATGACAGACCTTCAGGGTGGAA 480

Qy 1181 AGATATCACTACATACAGCCACTCACTAATGAAAGAGATGGATTCCTCCCACTAATATATTCAT 1240

Db 481 AATATCTAGAGACAGTCAGATATAAAGAAAGATGAGATTTCCCAATGACATCTCCT 540

Qy 1241 TGCTCAAGTTAAAGCCAGTGGATGGTCACTGTGCTCTAGAAATCCAAATACGTGAAGACTG 1300

Db 541 TCTCTAAGTTAAAGCCAGTGGTGGTCACTGTGCTCTGGAATCCAGATATGTGGAAGACTG 600

Qy 1301 TGTGCTTGCGCTGATGGGCTCTTCCCTCTGGGAGTGGATGCCACATCTGTGGCTGGGGGTG 1360

Db 601 TATGTTTGGACAGCGACCCCTTTGCCCTCTGGACGTGAAGGACACATCTGTGACATGGGTG 660

Qy 1361 TTACAGAAACAGG 1373

Db 661 GTACAGAAACAGG 673

BASE COUNT	188 a	221 c	188 g	154 t
ORIGIN				

	Query Match	25.9%	Score 435.4	DB 13	Length 751
	Best Local Similarity	78.0%	Pred. No. 7.2e-113		
	Matches 589	Conservative 0	Mismatches 156	Indels 10	Gaps 5
QY	4	TTTGGCAGATCTCGATCCATCCATTTTTCGCTGTAATGCTCTCGTGGGAAGACACG	-	62	
Db	2	TTTTCAGAGATTTTGATGTGTCTCCGCTCCGCTGCTCAATCCGCTCGATGGGAAGTCAGT	G	61	
QY	63	CTGTGGTTCCTCCGATGCTCTTATTGGAACCTGCACCACTGCACCCCTGACCA	122		
Db	62	CATTTGGCTCTACGATGCTCTTATTCGAGCCCGACACCAATTTGACCCCGGATGA	121		

```

OY 123 GTATGATACAGTACGAGATTATTAATCAGAAAGAACACCAGTACGACATTACCA 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 CTATATACAGTATGAGCAGTCCAGCCAGACAGAACGCCAGTGTCCAGCAGACCAC 181
OY 183 TGTGAGAACTCTGATGCTGACTACACTGAGACCAAGCTGATCCATGCCAGCCAAACC 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 CCGTAGAACCCCGAGTGTACT-----ATGAAGACGATGATCATCCAGTCCACCC 235
OY 243 CTGTAACACAGGAGGAGTGCCTGCTGCATGGAGACACCTTCACATGACGCTGCGC 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 CTGGAACACAGGAGGAGTGCCTGCTGCATGGAGACACCTTCACATGACGCTGCGC 295
OY 303 TCCCTTCTGGGAATTAAGTGCAGAAAGTGCAGAAATPAGCTCAGACAGAACCCATGTGG 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 296 CCCCCTCTCGGGAGCCGCTGCAGACAGCAAAACAGTGTCAAGGACAAACCCATGTGT 355
OY 363 CCGGGGCAATGTCTCATTTACACAGCTCCCTACTACCTGCTGTCTGTAACACCC 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 356 CCATGGTGAATGCTCATTTACCCAGAACACCCCTACTACCTGCTGTCTGTAACATGCC 415
OY 423 TTACACAGTCCAGCTGCTCCCAAGTGTCTGTAAGCAGGACCAACCCCTGCCAGAA 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 416 TTACACAGGACCCGAGCTGCTCCAAAGTGTCTGCGGATGAGCCAAACCCCTGCCAGAA 475
OY 483 TGGGGCTACTGCTGCTCCGACATACGGGAGATCCCAAGTTCACCTGCTGCCAGACA 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 476 TGGGGAGTGTGTTCCCGACACAGCAGAGATCCAGTTTACTGTGCTGTCCAGACCA 535
OY 543 GTTCAAGGGAATTTCTGGAATAGTTTCTGATGACTGCTATGTTGGGATGCTACTC 602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 536 GTATTAAGGGAATTTCTGGAATAGTTTCTGATGACTGCTATGTTGGGATGCTACTC 595
OY 603 TTACACAGGGAATTAAGTATGAGCAGTCAACAGTACGCTGCTTACTGAG-NACTGCC 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 596 TTACCGAGGCAAGTATGATGAGCAGTCAACAGCAGACCAACCCATGCTTACTGTAACACTGCC 655
OY 662 ACCTGCTCTTGCAGGAAATTAACAACATGTTTATGAGAGTGTGTAACCCCATGGAATGG 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 656 ACCCTCTCTTGCAGGAAATTAACAACATGTTTATGAGAGTGTGTAACCCCATGGAATGG-TC 714
OY 722 GGAACACAAATTTCTGCAGAAACCCAGATGGCGAG 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 715 GAGAGACAACTTCTG-AGAAACCCAGATGGAGAC 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
BM508620 570 bp mRNA linear EST 12-MAR-2002
LOCUS 1137H03.y1 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus
DEFINITION musculus cDNA clone IMAGE:5944156.5' similar to TR:Q14520 Q14520
GCF ACTIVATOR LIKE PROTEIN. ; mRNA sequence.
ACCESSION BM508620
VERSION BM508620.1 GI:18679763
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus;
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
1 (bases 1 to 570)
REFERENCE
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishke,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., McCann,R., Cole,R., Tsagaris,I., Williams,T.,
Jackson,T., and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: 1137H03.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138

```

```

Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For Information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
MGI:2007028 This sequence now available from the IMAGE consortium,
for clone orders contact: info@image.llnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 435.

FEATURES
    source
        1..570
            /organism="Mus musculus"
            /strain="TGR"
            /db_xref="taxon:10090"
            /clone="IMAGE:5944156"
            /clone_1fb="Melton Normalized Mixed Mouse Pancreas 1
            N1-MMS1"
            /sex="Both for embryonic & newborn, male for adult and
            adult Islet."
            /dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
            adult, mixed"
            /lab_host="DH10B"
            /note="Vector: pSPORT1; Site1: Not I; Site2: Sal I; Five
            libraries representing E10.5/12.5 pancreatic bud, E16.5
            pancreas, newborn pancreas, adult pancreas, and adult
            islets of langerhans were separately constructed using
            Superscript Plasmid Library Kit (Life Technologies). cDNA
            was made by oligo-dT priming and size selected by column
            fractionation. Libraries were amplified once on solid
            support and plasmid DNA from each library was prepared
            and mixed in equal amounts. The mixed library DNA was
            normalized by method #4 from Bonaldo, Lennon, and Soares
            1996 Genome Research 6:791-806. 0.5 microgram
            single-stranded mixed library plasmid DNA was mixed with
            5 micrograms PCR product representing mixed library
            inserts and hybridized to an EcoT of 6. Single-stranded
            (unhybridized) plasmids were isolated by hydroxyapatite
            chromatography and used to make this library."

BASE COUNT 162 a 134 c 152 g 122 t
ORIGIN
Query Match 25 7%; Score 432; DB 13; Length 570;
Best Local Similarity 85.0%; Pred. No. 5.8e-112;
Matches 483; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

OY 1108 CATCTAAGCTGTGCTTGGGAGCAGACCTGTAAGAAAGAAATTTTCATGACAGAC 1167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3 CATCTAAGCTGTGCTTGGGAGCAGACCTGTAAGAAAGAAATTTTCATGACAGAC 62
OY 1168 TTTAGGCTGCAGAAATTAAGTATGAGCAGTCAATGAAAGAGATGGATTTCCCGAC 1227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 TTCAGGCTGGAANAATAGTATGAGCAGTCAATGAAAGAGATGGATTTCCCGAC 122
OY 1228 AATGATATTGCAATTCCTCAAGTTAAAGCAGTGGATGCTACCTGCTTGAATCCAA 1287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 AATGATATTGCTTCTCAAGTTAAAGCAGTGGATGCTACCTGCTTGAATCCAA 182
OY 1288 TAGCTAAGACGTGTGCTTGGCTGATGAGGCTCTTCCCTGGGAGGTGCGACATC 1347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 183 TAGTATTAAGCTGTATGTTTCCAGGAGCCCTTCTCTGGAAGTGAATGCCATC 242
OY 1348 TCTGGCTGGGTGTATACGAAGAAAGGCTCCCGACCTCTGGATGCCAAAGTC 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 243 TCTGGCTGGGTGTATACGAAGAAAGGCTCCCGACCTCTGGATGCCAAAGTC 302
OY 1408 AAGCTATTTGCCAACACTTTGTGAACCTCCGCAACTCTATGACCAACATGATGATGC 1467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 303 AAGCTATTTGTACCTTTTGTGACCTCCGCAACTCTATGACCAACATGATGATGC 362
OY 1468 AGTATGATCTGTGACGAAGAAATCTTACGAACCTGGGCAAGACACCTGGAGGTGACTCT 1527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT 202 a 234 c 194 g 161 t
ORIGIN

Query Match 23.6%; Score 396.4; DB 13; Length 791;
Best Local Similarity 75.6%; Pred. No. 1e-101;

Matches 520; Conservative 0; Mismatches 161; Indels 7; Gaps 2;

OY 1 ATGTTGCGAGGATGCTGATCTCCATGTTCTGCTTAATAGGCTCTGGTGGAAACACA 60
DB 23 ATATTGTGCGAGGATGCTGATCTCCGTTCTCTCTATATCGCTGTGGGAAATCA 82
OY 61 GCGTGTGCTGCTCCCTGATGCTTTATGAAAGCTGACCCAGACAGCTGACCTGAC 120
DB 83 GTCATGGGCTCTCACTGATGCTTCATTTGCCCCAGACCCAGAAATGGACCCCGCAT 142
OY 121 CAGATGATTAACGCTAGCAGATTATATCAGAGACACACAGTACAGACTTACC 180
DB 143 GACTATTACTAAGCTATGAGCAGTCCAGCCAGACGAAACCCAGTGTACAGCAGACC 202
OY 181 CAGTGTAGAAATCTGACTGTACTACATGAGACCAAGCTGATCCATGCCACCCAC 240
DB 203 ACCCTGAGAACCCCGACTGTACT-----ATGAAAGAGATGATTCATGCCATGCCAAC 256
OY 241 CCGGTGGAACAGGTTGGGAGCTGCTGTCATGGAGACCTTACATGCACTGCTGCTG 300
DB 257 CCCTGTGAACAGGCGGAGAGCTGATATCATAGAGGGATATCTTGAATGCAAGTGCACA 316
OY 301 GCTCTCTTCTGAGGAATTAAGTGTGCAAAAGTCAAAATACGTGCAAGACACACCATGT 360
DB 317 GCCCCTTCTCGGGAGGCGCGTGCAGACATGCAAAAACAAAGTGCAGAACACCATGT 376
OY 361 GCGCGGGGCAATGCTATTAACCCAGAGTCTCTACTACGCTGCTGTGTAAACAC 420
DB 377 GTCCATGAGGATGTGCTCATTAACCCAGAACACCTCTACTACGCTGCTGTGAAATAC 436
OY 421 CTTTAAACAGATCCCACTGCTCCCAAGTGCTGTGTGTCAGAGGCAACCCCTGCGAC 480
DB 437 CTTTAAACAGGAGCACTGCTCCCAAGTGCTCCGCGATGAGGCAACCCCTGCGACA 496
OY 481 AATGGGAGTACTGCTCCGCGATTAAGCGAGATTCACATTCCTGCTGCTGCGGAC 540
DB 497 ATGGCGGAGTCTGTGTCCGACACAGACGAGATTCACATTCCTGCTGCTGCGGAC 556
OY 541 CAGTTAAGGGAATTTCTGTGA-ATAGTGTGTATGACGTCTATGTTGGCGATGGCTA 599
DB 557 CAGTATTAAGGGAATTTCTGTGAACATAGTCCGCGAGCTGTATGTGCGTGTATGCTA 616
OY 600 CTCTTACGAGGGAATTAATAGACAGTCAACAGCATGCTGCTTCTTCTGGAACCTG 659
DB 617 CTCTTACGAGGGAATTAATAGACAGTTCACAGAACCATCTTCTTCTGGAACCTG 676
OY 660 CCACCTCTCTCTGCGAGGGAATTAACAC 687
DB 677 CCAGCTCTCTCTGCGAGGGAATTAAC 704

RESULT 9
AA237499 546 bp mRNA linear EST 03-MAR-1997
LOCUS mxi0b10.t1 Soares mouse NML Mus musculus cdna clone IMAGE:679771.5
DEFINITION similar to TR:G1345398 G1345398 HGF ACTIVATOR LIKE PROTEIN. ; mRNA
sequence.
ACCESSION AA237499
VERSION AA237499.1 GI:1861520
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 546)

AUTHORS

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisell, S., Kucada, F., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

JOURNAL

The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project

COMMENT

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:419475
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 405.
location/Qualifiers

FEATURES

source

1..546

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:679771"

/clone_lib="Soares mouse NML"

/tissue_type="Liver"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15'
TGTACCAATCTGAGTGGAGCGCGCGGAACTTTTCTTTTCTTTT 3';
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."

BASE COUNT 141 a 142 c 151 g 112 t
ORIGIN

Query Match 23.4%; Score 394.2; DB 9; Length 546;
Best Local Similarity 82.9%; Pred. No. 3.5e-101;

Matches 450; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

OY 868 GAGCCATCAACAGCTTCCGGGTTTGACTCTCTGTGAAAGACTGATAGCAGAGG 927
DB 4 GAGCCTGTGATGAGAGCTGCGGAGGTTGAGTCTGAGGGAAGACGAGGATGCTGAAAC 63
OY 928 AAGATCAAGACATCTATGAGGCTTTAAGAGACAGCGGCGGCAAGCCATGGGAGCG 987
DB 64 GCAGTCAAGCTATCTAGGAGGCTTTAAGAGACAGAGCAAGACCGTGAGAGG 123
OY 988 TCCCTCAGTCTCTGCTGCTGTGACCATTCATGCCCCAGGCGCACTTGTGTGGG 1047
DB 124 TCCCTCAGACCTCCTACTGCGGTTGACCATTCATGCCCCCAAGGCCACTTGTGTGGG 183
OY 1048 GCGGTGATGCCACCCCTGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1107
DB 184 GCGGTGATGCCACCCCTGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243
OY 1108 CATCTAAAGTGTGCTAGGAGGACGACCTGAAAGAAAGAAATTTCTGAGCAGAGC 1167
DB 244 CATCTAAAGTGTGCTAGGAGGACGACCTGAAAGAAAGAAATTTCTGAGCAGAGC 303
OY 1168 TTTTAGGTGCGAAGATTAATCAAGTACAGCCATCAATGAAGAGATGATGTCCTCCAC 1227
DB 304 TTTAGGTGCGAAGATTAATCAAGTACAGCCATCAATGAAGAGATGATGTCCTCCAC 363
OY 1228 AATGATTAATGCTGCTCAAGTTAAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1287
DB 364 AATGATTAATGCTGCTCAAGTTAAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 423
OY 1288 TACGTGAAGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1347

Accession	Sequence	Position
Db	TATGTAAAGACTGATGTTTGTCCCAAGCACCCTTTCCCTCTGTGAACCTGAGTGCACATC	483
Db	424	
QY	TCGCGCTGGGGTGTATTACAGAAACAGCAAAAGAGGTGCCCGCAGCTCCTGTGATGCCAAAGTC	1407
Db	484	
QY	TCGCGCTGGGGGTGTACAGAAACCAAGGAGGGTTCGCGCAGCTCCTGTGATGCCAAAGTC	543
QY	1408	
Db	544	

RESULT	10
A6601564	
LOCUS	
DEFINITION	A6601564 Bos taurus kidney fetus Bos taurus CDNA clone EST 27-NOV-2001 5' mRNA linear
ACCESSION	A6601564
VERSION	A6601564
KEYWORDS	EST.
SOURCE	COW.
ORGANISM	Bos taurus

REFERENCE
AUTHORS
1 (bases 1 to 608)
Takasuga, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H.

TITLE	Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
JOURNAL	Nucleic Acids Res. 29 (22), E108 (2001)
MEDLINE	21570554

COMMENT

Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel.: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@ccoca.ocn.ne.jp

Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES	source
Location/Qualifiers	
1. .608	
/organism="Bos taurus"	
/db_xref="taxon:9913"	
/clone="PIKT005A04"	
/clone_1ib="Bos taurus kidney fetus"	
/tissue_type="kidney"	
/dev_stage="fetus"	
/lab_host="DH10B"	
/note=Vector: pZLL; Site.1: SalI; Site.2: NotI; Poly A	
was deleted from a NotI site"	
148 a	171 c
151 g	135 t
3	others

	Query Match	23.3%	Score 392.6	DB 10:	Length 608;
	Best Local Similarity	81.9%	Pred. No. 1.1e-100;		
	Matches 465;	Conservative 0;	Mismatches 97;	Indels 6;	Gaps 1.
Oy	1	ATGTTTGCCAGSAGTGTGCATTCACATGTTCTGCTGTAAITGGCTCTGTGGTAAGAACA	60		
Dd	47	ATGTTTTGCCAAGATGTGCATTCACATGTTCTGCTGTAAITGGCTCTGTGGTAAGAACA	106		
Oy	61	GCCTTGGSGTTCCTCCGTGATGTCTTTATTGGAAACCCTGGAACCCACTGGACCCTCTGAC	120		
Dd	107	GCCTTGGSGCTCTCCCTCGCTGTCTTTTTCACAGAGCCGGAGCCCAATTGGACTCTTGAC	166		
Oy	121	CAGTATGATTACAGTACGAGATTAATACTCGASAGGAACACCCATGTAGCACACTTACC	180		
Dd	167	CAGTATGAGATTACGACGAAGATTAATAACCGASAGGAACCCCATGTAGCACACTTACC	226		
Oy	181	CATGCTGAGATCTCTGACTGTGACTACACTGAGGAGCACAAGCTGATATCCATGCCACGCCAAC	240		

Db	227	TATTCGTACACACCTGACTGTGTACTACGAAAGAGA-----TACCCATGCGCTGCCAC	280
QY	241	CCCTGTGAACACGGTGGGGACTCCTCGTCATGGGAGCACCTTCATCATGCACGCTGGCTG	300
Db	281	CCCTGTACACATGGTGGGGACTCCTCGTCAGTGGGGGCCACCTTCACGTCGGCGCTGGCCG	340
QY	301	GCCTCTTCTCTGGGAATAAGTGTGAGAAAGTGCAGAAATACGTGCAGGACACCACTGT	360
Db	341	GACCTTTCTCTGGAAACAGGTGTCTCAGATGTGTGAAACAGACGCAAAACACACCTGT	400
QY	361	GGCGCGGGGCCAATGTCTCATTAACGAGGTCTCCCTACTACCGCTGTGTCTGTAAACAC	420
Db	401	GGCGGGGAGACTGTCTCATTAACGAGGTCTCCCTTACACCGCTGTGTCTGTAAACAC	460
QY	421	CCCTTCACAGGTGCCAGCTGTGTGCCACAGGGTTCCTGTATGCAAGGTCAAAACCCCGGCAG	480
Db	461	CCCTTCACAGGGGTGTAGACGTCTTCMCAGAGGGTCTGTGTGCAAGGTCAAAATCCTGTMCNA	520
QY	481	AATGGGGGTACTCTCTCCCGGCATAAGCGGAGATCCAAATTCACCTGTGGCCCTGTGCCGAC	540
Db	521	AACGGGTGCACCTCTCTCCCGGCAGANAGCGGAGGTCCAATTCACCTGTGGCCCTGTGCAC	580
QY	541	CACCTTCANAGGGGAATTCGTGCAATAG	568
Db	581	CAGTTCANAGGGGAAGCTTTTGCAATAG	608

RESULT	11				
LOCUS	Bf780971				
DEFINITION	Bf780971	902 bp	mRNA	linear	EST 12-JAN-2001
ACCESSION	602105495F1	NCI-CGAP_Kid1A	Mus musculus	CDNA clone	IMAGE:4223710
VERSION	Bf780971	5', mRNA sequence.			
KEYWORDS	Bf780971.1	GI:12086004			
SOURCE	house mouse.				
ORGANISM	Mus musculus				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 902)	EXAIRYOTA; METAZOA; CHORDATA; CRANIALATA; VERTEBRATA; EUCELOSTOMI; MAMMALIA; EUTHERIA; RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.			
	NIH-MGC	http://mgc.nci.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	
		Unpublished (1999)		
	Contact: Robert Stransberg, Ph.D.	Email: cgaps-r@mail.nih.gov		
	Tissue Procurement: Jeffrey E. Green, M.D.	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:	http://image.llnl.gov		
	Plate: LLAM5612	row: 1	column: 23	
	High quality sequence stop: 17.			

```

source
1. .902
   /organism="Mus musculus"
   /strain="FVB/N"
   /db_xref="taxon:10090"
   /clone_image:423710"
   /clone_lib="NCI_CGAP_kid14"
   /lab_host="DH10B (T1 phage-resistant)"
   /note="Organ: kidney; Vector: pCMV-SPOrt6; Site:1; Motif:
   Site:2; Salt; Cloned unidirectionally. Primer: Oligo dT.
   Average insert size 1.75 kb. Constructed by Life
   Technologies. Note: this is a NCI_CGAP Library. "
BASE COUNT      223 a      264 c      254 g      161 t
ORIGIN

Query Match      23 2%; Score 390.4; DB 12; Length 902;
Best Local Similarity 73.18; Pctd No. 5.5e-100;
Matches 532; Conservative 0; Mismatches 186; Indels 10; Gaps 2

```

```

OY 1 ATGTTGCCAGATGTCGATGTCATGCTGCTGTTATATGAGCTGTGGGAAAGACA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 52 ATATTTCAGAGATGTTGTTGTTCCGTCTGCTGTTATGCTGCTGTTGGGAGATCA 111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 GCCTGAGGTTCTCCCTGATGCTTTATTTGAAACCTGAGCCAGACATGAGACCCCTGAC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 112 GTCATTGGGCTCTCATGATGCTTCATTGCGCCCAAGACCAAGATTTGGACCCCGAT 171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 121 CAGTATGATGATGCTAGCAGATTTATATCAGAGAGAGACCAAGCTAGCACTTAC 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 172 GACTATTACTACAGCTATGAGCATGCCAGCCAGACAGAACCCAGTGTCCAGACAC 231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 181 CAGTCTGAGATCTGATGCTGATCTACTGAGACCAAGCTGATCCATGCCAGCCAC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 232 ACCCTGAGAACCCGACTGCTACT-----ATGAAACAGATGATCATGCTCAGTCCAC 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 241 CCCTGTACACAGGCTGGGAGCTGCTGCTCCATGAGGAGCAACCTTACATGACGCTGCTG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 286 CCCTGTACACAGGCTGGGAGCTGCTGCTCCATGAGGAGCAACCTTACATGACGCTGCTG 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 301 GCTCTTTCTCTGGAATAGTGTGAGAAAGTGCAGAAATACGTGCAAGGACAAACCATGT 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 346 GCCCCTTCTCGGGAGCGGTGCGAGACTGCACAAACAAAGTGCAGAGACAAACCATGT 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 361 GCGCGGGGCAATGCTCATTTACCCAGAGTCTCCCTACTACGCTGCTGCTGTAAACAC 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 406 GTCATGCTGATGCTCTCATTTACCCAGAGCAACCCCTACTACGCTGCTGCTGTAAATAC 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 421 CCTTACACAGATGCCAGCTCTCCCAAGTGGTTCCTATATGACAGGCCAAACCCCTGAC 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 466 CCTTACACAGGAGCAACAGCTCTCCCAAGTGGTTCCTATATGACAGGCCAAACCCCTGAC 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 481 AATGGGCTACCTGCTCCCGGATGAGGAGATCCAGATTCACCTGCTGCTGCTGCTGAC 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 526 AATGGGCTACCTGCTCCCGGATGAGGAGATCCAGATTCACCTGCTGCTGCTGCTGAC 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 541 CAGTTCAAGGGGAAATTTCTGTAATAGTTCTGATGATGCTGATGTTGGCGATGCTAC 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 586 CAGTATTAAGGGGAAATTTCTGTAATAGTTCTGATGATGCTGATGTTGGCGATGCTAC 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 601 TCTTACGAGGAGAAATGTAATGAGACATCAACAGATGCTGCTTACTGAGAACTCC 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 646 TCTTACGAGGAGAAATGTAATGAGACATCAACAGATGCTGCTTACTGAGAACTCC 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 661 CAGTCTCTCTGAGAGAAATTTACACATGTTTATGAGAGATGCTGAAACCCATGGATTT 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 706 CTCTCTC---TGGCGGAGATTTCTAGCGGTATGAGGAGGTGCGGAGACCCCGGAACG 761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 721 GGGGACACA 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 762 CGGACACA 769
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://Image.lnl.gov
Plate: L1236
High quality sequence start: 2
High quality sequence stop: 682.
Location/Qualifiers
1..682
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5098442"
/clone_1lb="NCI CGAP L19"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 Kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT
166 a 206 c 169 g 141 t
ORIGIN
Query Match 22.9%; Score 386; DB 13; Length 682;
Best Local Similarity 77.0%; Pred No. 8.6e-99;
Matches 511; Conservative 0; Mismatches 145; Indels 8; Gaps 3;
OY 1 ATGTTGCCAGATGTCGATGTCATGCTGCTGTTATATGAGCTGTGGGAAAGACA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 24 ATATTTCAGAGATGTTGTTGTTCCGTCTGCTGTTATGCTGCTGTTGGGAGATCA 83
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 GCCTGAGGTTCTCCCTGATGCTTTATTTGAAACCTGAGCCAGACATGAGACCCCTGAC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 84 GTCATTGGGCTCTCATGATGCTTCATTGCGCCCAAGACCAAGATTTGGACCCCGAT 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 121 CAGTATGATGATGCTAGCAGATTTATATCAGAGAGAGACCAAGCTAGCACTTAC 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 144 GACTATTACTACAGTATGAGCATCCAGCCAGAGAGACCCAGTGTACGAGAC 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 181 CAGTCTGAGATCTGATGCTGATCTACTGAGACCAAGCTGATCCATGCCAGCCAC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 204 ACCCTGAGAACCCGACTGCTACT-----ATGAAAGAGATGATCCATGCCAGTCCAC 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 241 CCCTGTGAACAGGCTGGGAGCTGCTGCTCATGAGAGACCTTACATGACCTGCTG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 258 CCCTGTGAACAGGCTGGGAGCTGATCATCAGAGGAGATCCCTCAGTGTGAGCTGCCA 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 301 GCTCTTTCTCTGGAATAGTGTGAGAAAGTGCAGAAATACGTGCAAGGACAAACCATGT 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 318 GCCCCTTCTCGGGAGCGGTGCGAGACTGCACAAACCAAGTGCAGAGACCCCGATGT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 361 GCGCGGGGCAATGCTCTCATTTACCCAGAGTCTCCCTACTACGCTGCTGTAAACAC 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 378 GTCATGATGATGATGCTCTCATTTACCCAGAAACCAACCTTACTACGCTGCTGCAATAC 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 421 CCTTACACAGGCTCCAGGCTGCTCCAGATGTTCTGATGAGAGCCAAACCCCTGAC 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 438 CCTTACACAGGAGACAGACTGCTCCAAAGTCTTCCGGCATGTGAGGCC-AAACCTGTCCAG 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 481 AATGGGCTACCTGCTCCCGGATGAGGAGATCCAGTTCACCTGCTGCTGCCAG 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 497 AATGGGAGATGCTGTTCCCGACAGACAGAGATCCAGTTCACCTGCTGCTGCCAG 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 541 CAGTATTAAGGGGAAATTTCTGTAATAGTTCTGATGATGCTGATGTTGGCGATGCTAC 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 557 CAGTATTAAGGGGAAATTTCTGTAATAGTTCTGATGATGCTGATGTTGGCGATGCTAC 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 601 TCTTACCGAGGAGAAATGAAT-AGGACAGTCAACGACATGCTGCTTACTGGAATC 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 617 TCTTACCGAGGAGAAATGAATGAATGAACAGTCAACGAGACCATGCTTACTGGAATC 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 660 CCAC 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      677 CCAC 680
|||||
RESULT 13
LOCUS   BF785781
DEFINITION 929 bp mRNA linear EST 12-JAN-2001
602112402F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4240465
5', mRNA sequence.
ACCESSION BF785781
VERSION   BF785781.1 GI:12090817
KEYWORDS EST.
SOURCE   house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 929)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM9855 row: d column: 02
High quality sequence stop: 625.
Location/Qualifiers
1. 929
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1db="IMAGE:4240465"
/clone_1lb="NCI_CGAP_Kid14"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site:1; NotI;
Site:2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. 1"

BASE COUNT      238 a      269 c      237 g      185 t
ORIGIN
Query Match      22.9%; Score 384.6; DB 12; Length 929;
Best Local Similarity 79.8%; Pred. No. 2.5e-98;
Matches 478; Conservative 0; Mismatches 119; Indels 2; Gaps 2;

QY 372 ATGTCCTATTCCACAGTCCCTCTACTACCGCTGTCTGTAAACACCTTACACAG 431
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19 ATGCTCTATTACCCAGAGACCCCTACTACCGCTGTCTGTAAACACCTTACACAG 78
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 432 TCCACAGCTGCTCCCAAGTGTCTGTATGACGAGCAACCCCTGCCAGATGGGGCTAC 491
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 ACCAGACTGCTCCAAAGTGTCTGTATGACGAGCAACCCCTGCCAGATGGGGAGT 138
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 492 CTGCTCCCGGATTAAGGAGATCCCAAGTTCACCTGTGCTGCTCCGACCAAGTTAAAGG 551
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139 CTGTTCCGACACGACGAGGATCCAGGTTTACCTGTGCTGCTCCAGACCAAGTTAAAGG 198
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 552 GAAATTCGTGAAATAGTCTGATGACTGCTATGTTGGGATGGCTTACCTTACCGAG 611
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 GAAATTCGTGAAATAGTCTGATGACTGCTATGTTGGGATGGCTTACCTTACCGAG 258
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 612 GAAATGAATGAGACATCAACGACATGCTGCTTACTGCAACTCCACCTCTCTT 671
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 CAAATGATGATGAGACATCAACGACATGCTGCTTACTGCAACTCCACCTCTCTT 318
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 672 GCAGAGAAATTAACATGTTATGAGAGATGTAAGAACCATGAGTGGGAGAACAA 731
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 319 GCAGAGACTTATACATGTTATGAGAGATGTAAGAACCATGAGTGGGAGAACAA 738
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 732 TTTCGACAAACCCACATCGGACGAAAGCCCTGGTCTTTATTAAATTACCAATGA 791
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 379 CTTCGTGAGAAACCCAGATGGAGACCAAAACCCCTGTGTTTCGTAAGAGTGAACATGA 438
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 792 CAAGGTGAATGGGAATACGTGTGATGCTCAGCCTGTCAAGCCAGGAGCTTGCCTACCC 851
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 GAAGGTGAATGGGAATACGTGTGATGCTCAGCCTGTCAAGCCAGGAGCTTGCCTACCC 498
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 852 AGAGAAAGCCCACTGAGCCATCAACCAAGCTTCCGGGGTTTGACTCTGTGGAAGAC 911
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 AGTGAAGCCCTTCTGAGAGCCTGTGATGGAGC-TGCAGGCTTCGAGTCTCGGGAGAC 557
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 912 TGAGATGACAGAGAGAAATCAAGATCTATGAGGCTTTAAGACACGAGGGGCA 970
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 558 GAGAGTACGTGAACACGACGATCAAGCTATCTACG-GGCTTTAAGACACAGCAGGCA 615
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
LOCUS   BF384535
DEFINITION 969 bp mRNA linear EST 27-NOV-2000
602046804F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4196376 5',
mRNA sequence.
ACCESSION BF384535
VERSION   BF384535.1 GI:11365840
KEYWORDS EST.
SOURCE   house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 969)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM9533 row: g column: 01
High quality sequence stop: 692.
Location/Qualifiers
1. 969
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1db="IMAGE:4196376"
/clone_1lb="NCI_CGAP_L19"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site:1; NotI;
Site:2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. 1"

BASE COUNT      258 a      278 c      260 g      172 t      1 others
ORIGIN
Query Match      22.1%; Score 372.4; DB 12; Length 969;
Best Local Similarity 76.5%; Pred. No. 7.9e-95;
Matches 484; Conservative 0; Mismatches 142; Indels 7; Gaps 2;

QY 1 ATCTTTGCCAGATATCTGATTCATGCTTGTGTTAATGAGCTGTGGGAAAGACA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 ATATTTGTAGATGTTGGTGTTCCTGTCCTGCTGCTAATGCGCTGTGGGGAAGTCA 123
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 GCCTGTGGTCTCTCCGATGATGCTTTATGGAAGACCTGGACCCGACACTGACCTGAC 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 GTCATGGGCTCTCTCCGATGATGCTTTATGGAAGACCTGGACCCGACACTGACCTGAC 183
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 CAGTATGATTACAGCTACGAGGATTTAATCAGAGAAAGAACCAACCACTAGCATTACC 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```
Db 184 GACTATTACTATAGCTATGACAGTCCAGCCCGAGGAGGCCAGGTGTCACGACACC 243
QY 181 CANGTCAGAAATCTGACTGTGACTTACACCTGAGGACCAAGCTGATTCATGCGACGCCAAC 240
Db 244 ACCCGTAGAAGACCCCGACTGTGACT-----ATGAAGACGATGATCCATGCGACGTCCAAC 297
QY 241 CCTGTGAAACAGGTGGGAGTGTGCTGTCATGAGGACCTTCACATGACGCTGCTG 300
Db 298 CCTGTGAAACAGGTGGGAGTGTGCTGTCATGAGGACCTTCACATGAGGACCTGCTG 357
QY 301 GCTGCTTCTGCTGGAATTAAGTGTGAGAAAGTCAAAATACGTGCAAGACACCATGT 360
Db 358 GCCCGCTTCTGCGGAGGCGGTGCGCAAGCTGCACAAACAACTGCAAGACACCATGT 417
QY 361 GGGCGGGGCGCATGTCTCATTAACCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 418 GTCCATGTGTGATGCTGCTCATTAACCCAGACACCATCTACTACCGCTGCTGCTGCTGCT 477
QY 421 CTTTACACAGGTCCAGCTGCTCCCAAGTGTCTGTATGAGGCGCAACCCCTGCGCAG 480
Db 478 CTTTACACAGGTCCAGCTGCTCCCAAGTGTCTGTATGAGGCGCAACCCCTGCGCAG 537
QY 481 AATGGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 538 AATGGGCGAGTCTGTTCCCGACACAGACGAGATCCAGGTTTACTGCTGCTGCTGCTGCTG 597
QY 541 CAGTTACAGGGAATTTCTGGAATAGTTGATGATGCTGCTGATGTTGGCGATGCTTAC 600
Db 598 CAGTTACAGGGAATTTCTGGAATAGTTGATGATGCTGCTGATGTTGGCGATGCTTAC 657
QY 601 TCTTACCGAGGGAATGATAGACAGTCAC 633
Db 658 TCTTACCGA-GGCAAGTGTAGTACAGTCAC 689

RESULT 15
AM475402 654 bp mRNA linear EST 24-FEB-2000
LOCUS IMAGE:2503549.5 similar to TR:Q14520 Q14520 HGF ACTIVATOR LIKE
DEFINITION PROTEIN. ; mRNA sequence.
ACCESSION AM475402
VERSION AM475402.1 GI:7045508
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 654)
AUTHORS Maria,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T.,
Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Riltter
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterson,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
TITLE JOURNAL
COMMENT Contact: Maria M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LBNL; contact the
IMAGE Consortium (Info@image.lbnl.gov) for further information.
MGI:1018401
Seq primer: custom primer used
High quality sequence stop: 482.
FEATURES
Source 1..654
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
```

```
/clone="IMAGE:2503549"
/clone_lib="Sugano mouse kidney mklA"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pME18S-FL3; Site_1: DraIII
(cactgtgtg); Site_2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTATTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TGTGGCTTACTGG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGTGCTCTAAAGCTGCG and 3' end
primer CGACCTGCACTGCGACCA."
BASE COUNT 156 a 193 c 174 g 131 t
ORIGIN
Query Match 22.0%; Score 370.6; DB 10; Length 654;
Best Local Similarity 74.5%; Pred. No. 2,le-94;
Matches 482; Conservative 0; Mismatches 159; Indels 6; Gaps 1;
QY 1 ATGTTGCCAGAGTGTGATCTTCATGTCTTCTGCTGTTAATGCTGTGGTGGGAAAGACA 60
Db 13 ATATTGTGACAGATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 72
QY 61 GCTGTGGTCTTCTCCCTGATGCTTATTTGGAAGAACCTGAGACCGACAGTCCGCTGAC 120
Db 73 GTCAATTTGCTCTCAGTATGATGCTTCTTATTTGCGCCCGCCAGCCAGTATGAGACCCCAT 132
QY 121 CAGTATGATTTACCTACGAGGATTTATTCAGAAAGAACCCAGTATGACACTTAC 180
Db 133 GACTATTACTACAGCTATGAGCAGTCCAGCCAGACGAGAACCCAGTATGAGCAGAC 192
QY 181 CAGTGTGAGAAATCTGACTGTACTACACTGAGACCAAGCTGATCCAGCCGAC 240
Db 193 ACCCGTGTGAGAACCCCGACTGTGACT-----ATGAAGACGATGATCATCTCCAGGCAAC 246
QY 241 CCTGTGAAACAGGTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 247 CCTGTGAAACAGGTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
QY 301 GCTCTTCTCTGGAATTAAGTGTGAGAAAGTCAAAATTAAGTGTGAGACACACCATGT 360
Db 307 GCCCGCTTCTGCGGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
QY 361 GCGCGGGCGCAATGCTCATTAACCCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 367 GTCAATGTGATGCTCTATTAACCCAGAGACCCCTTAAGTGTGCTGCTGCTGCTGCTGCTG 426
QY 421 CCTTACAGAGTCTCCAGCTCTCCCAAGTGTCTGATGTCAGGACCAAAACCCCTGCGCAG 480
Db 427 CCTTACAGAGGACACAGACTCTCCCAAGTGTCTGATGTCAGGACCAAAACCCCTGCGCAG 486
QY 481 AATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 487 AATGGGCGAGTCTGTTCCCGACACACAGCAGATCCAGAGGTACTGCTGCTGCTGCTGCTG 546
QY 541 CAGTTCAAGGGGAAATTTCTGGAATTAAGTGTGATGATGTTGTCGATGCTGCTGCTGCTG 600
Db 547 CAGCTATAGGCGGAAATTTCTGGAATTAAGTGTGATGATGTTGTCGATGCTGCTGCTGCTG 606
QY 601 TCTTACCGAGGGAATTAAGTGTGATGATGTTGTCGATGCTGCTGCTGCTGCTGCTGCTGCTG 647
Db 607 TCTTACCGAGGGAATTAAGTGTGATGATGTTGTCGATGCTGCTGCTGCTGCTGCTGCTGCTG 653
```

Search completed: March 6, 2003, 19:37:34
Job time : 2184.5 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 15:27:01 ; Search time 65.5 Seconds

(Without alignments)
7879.954 Million cell updates/sec

Title: US-09-912-559-2

Perfect score: 1683

Sequence: 1 acgttcgccagatgtctga.....aaagtgaagtgcctctaa 1683

Scoring table: IDENTITY_NUC

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents,NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfileseq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114.4	6.8	2033	1	US-08-148-910-14
2	114.4	6.8	2033	1	US-08-148-910-14
3	93	5.5	970	1	US-08-148-910-3
4	93	5.5	970	1	US-08-148-910-3
5	72.2	4.3	329	1	US-08-148-910-2
6	72.2	4.3	329	1	US-08-148-910-2
7	72.2	4.3	329	1	US-08-148-910-13
8	72.2	4.3	329	1	US-08-148-910-13
9	72.2	4.3	1065	1	US-08-427-640-1
10	72.2	4.3	1065	1	US-08-427-640-1
11	72.2	4.3	1068	1	US-08-427-640-3
12	72.2	4.3	1137	4	US-09-553-498-9
13	72.2	4.3	1137	4	US-09-553-498-9
14	72.2	4.3	1314	2	US-08-811-949-48
15	72.2	4.3	1314	2	US-08-811-949-48
16	72.2	4.3	2457	6	US-08-885-795A-39
17	72.2	4.3	2457	6	US-08-885-795A-39
18	72.2	4.3	7360	1	US-08-286-740-1
19	70.6	4.2	7360	5	US-08-286-740-1
20	70.6	4.2	1068	1	US-08-811-949-60
21	70.6	4.2	1068	1	US-08-811-949-60
22	70.6	4.2	1068	2	US-08-427-640-7
23	70.6	4.2	1068	2	US-08-427-640-7
24	70.6	4.2	1068	2	US-08-811-949-44
25	70.6	4.2	1068	2	US-08-811-949-46
26	70.6	4.2	1068	2	US-08-811-949-52
27	70.6	4.2	1068	6	US-08-811-949-58

28	70.6	4.2	1163	4	US-09-410-882-5	Sequence 5, Appl
29	70.6	4.2	1170	2	US-08-811-949-64	Sequence 64, Appl
30	70.6	4.2	1170	2	US-08-811-949-66	Sequence 66, Appl
31	70.6	4.2	1314	2	US-08-811-949-50	Sequence 50, Appl
32	70.6	4.2	1314	2	US-08-811-949-54	Sequence 54, Appl
33	70.6	4.2	1314	2	US-08-811-949-56	Sequence 56, Appl
34	70.6	4.2	1419	2	US-08-811-949-62	Sequence 62, Appl
35	70.6	4.2	1738	6	US-08-811-949-62	Sequence 62, Appl
36	70.6	4.2	1848	3	US-08-814-412-10	Sequence 10, Appl
37	70.6	4.2	1974	2	US-08-811-949-38	Sequence 38, Appl
38	70.6	4.2	2101	2	US-08-811-949-42	Sequence 42, Appl
39	70.6	4.2	2162	1	US-08-119-512-3	Sequence 3, Appl
40	70.6	4.2	2162	1	US-08-488-015B-3	Sequence 25, Appl
41	70.6	4.2	2162	1	US-08-488-015B-25	Sequence 25, Appl
42	67.8	4.0	1724	6	US-09-518-046-3	Sequence 3, Appl
43	67.8	4.0	2497	6	US-09-518-046-3	Sequence 3, Appl
44	66.4	3.9	2544	4	US-08-254-922-1	Sequence 1, Appl
45	66	3.9	1233	1	US-08-254-922-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-148-910-14
Sequence 14, Application US/08148910
Patent No. 546593
GENERAL INFORMATION:
APPLICANT: Takeshi SHIMOMURA et al.
TITLE OF INVENTION: No. 546593el Protein and Gene Encoding Said Protein
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
City: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch,
MEDIUM TYPE: 500 Kb Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/148,910
FILING DATE: No. 546593ember 5, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: human
IMMEDIATE SOURCE:
LIBRARY: Pre-made Lambda phage Library,
LIBRARY: human liver(49, male) cDNA Library (Stratagene)
US-08-148-910-14

Query Match	Similarity	49.6%	Pred. No. 3.5e-24	DB 1	Length 2033:
Matches	605;	Conservative	0;	Mismatches 561;	Indels 54;
					Gaps 10;
QY	469	AACCCCTGCCAGATGAGGCGTACCTCTCCCGGACATTAAGCGGAGATCCAAATGACCTGT	528		
Db	742	AGCCCTTGCTCCCTACACGGGGGACCTCCACCTGATCGTGGCCACCGGACACCGTGTGT	801		
QY	529	GCCGTGCCGACCACTTTCAGGGGAAATTCGTGAATAGTTCTGATCA---CTGGCAT	585		
Db	802	GCCGTGCCACACGAGCTTGCGGAGCGCTCTGCAACATCAGCCTTATATAGCGCTTCT	861		
QY	586	GTTGGCGATGCTACTCTTACCGAGGAAATGAAATGATGACAGTCAACCGATCGCTGC	645		
Db	862	TTGGGGAGAGGACACTGGGTACCGCTGGCGCGACACCTCAGACCTCGGGCTCAGCTGC	921		
QY	646	CTTTACTGGAACTCCACACTCTCTTGGACGGAATTAACAATGTTTATGGAGATGCT	705		
Db	922	CTGGCCTGGAACTCGATCTGCTCTACAGAGAGCTGCACGCTGACCTCGTGGCGCCGC	981		
QY	706	GAAACCATGAGATTTGGGAGACAAATTTCTGCAGAAAACCCAGATGCGAGAAAACCC	765		
Db	982	GCCCTCTGCGCTGGGCCCCCAGTGCCTTACGCGGATCCGACAAATGAGAGAGGCC	1041		
QY	766	TGCTGTTTAAATTAAGTTTACCAATGACAGTGAAGAAATGGGAATGCTGATGCTCAGCC	825		
Db	1042	TGGTGCTAGCT---GTTGAAGSACAGCGGCTCTCTGAGAGTACGCGCTTGGAGGCT	1098		
QY	826	TGCTCAGCCCAAGAGTGGCTTACCACAGGAAAGCCCATAGCATCAACAGCTT	885		
Db	1099	TGGGAATCCCT-----CACAGATCCAACTGTCACCGGATCTCTGGACACCTGT	1149		
QY	886	CCGGGGTTTGACTCTCTGGAAGAAGTGAATAGACAGAGGAAGTCAAGAGATATAT	945		
Db	1150	CTGAGACCTAGCTCTCCCGGGGCCAGCGCTGGCGGACGAGGACAAAGAGAGCTTC	1209		
QY	946	GGAGGCTTTAAGACAGCGGGGSCAAGCACCATATGGAGGCTCTCATCTCTGCTG	1005		
Db	1210	CTCGGGCCAGTATATCGGGGCTCTCTCTGCTGCGCCGCTCGACCTCGT-----	1262		
QY	1006	CTCTGTACCATCTCCATGCCCCAGGGCCACTCTGTGTGGGGCGCTGATCCACCTCTG	1065		
Db	1263	--GCTGGCGGCATCTACATCGGGGACACCTTCTGCGCGGAGCCCTGGTCCACACTGG	1320		
QY	1066	TGGGTCTCAGCGCGCCGACCTGACACCGACATMAAAACAGACATCTAA-----AGCTG	1119		
Db	1321	TGGGTGTGTGCGCCGCCCACTCTTCTCCACAGGCCCCCCAGGAGACGCTTCCGTG	1380		
QY	1120	GTCCTAGGGGACAGGACCTGAAAGAAAGAAATTTTCATGACAGAGTCTTAGGGTGCAG	1179		
Db	1381	GTCCTGGGCGACACTTCTTCAACGCCACGACGACGACGTCGACGACGACCTGGCATGAG	1440		
QY	1180	AAGATTTTCAAGTCAACCCACTCATATGAAGAAGATGATGTTCCCAATAGATATGCA	1239		
Db	1441	AAGTATATCCGTACACCTCTGATCTGGGTTCACACCCAG---GACACAGACCTGTGTC	1497		
QY	1240	TTCCTCAGATTAAAGCCAGTGGATGGTCACTGCTCTAGATATCCAAATACGTGAAGCT	1299		
Db	1498	CTGATCCGCTGAAGAAAGAAAGGGACCGCTGTGTCACAGCTCCGATCTGTGCACGCC	1557		
QY	1300	GTTGTGTTGCTGCTGATG---GATCTTCTTCCCTGTGGAGAGTGGTGCATCTCTGGC	1353		
Db	1558	ATCTGCTGCTCCGACGCCGCGACACACTTCCCGCAGGACACAAATGCTGCGGCG	1617		
QY	1354	TGGGG-----TGTTCACAGAAACAGGAAAGGTCGCCGACGCTCTGTGACCAATGTC	1407		
Db	1618	TGGGGGCACTTGGATGAGAAAGTGAAGCGGCTACTCTAGCTCCCTCGGAGAGCCCTGGTCTC	1677		
QY	1408	AAGCTGATTTGACCACTTTGTGCAACTCCGACCACTCTATAGACACAGATGATGAC	1467		
Db	1678	CCCTGTGTGCGGACACCAAGTGCACAGCCTTAGGTCTCAGGCGCCGACATAGCCCC	1737		
QY	1468	AGTATGATCTGTGAGGAATCTTCAAGAAACCTGGGACAGACACCTGCAAGGTGATCT	1527		

```

Db      1738  AACATCTCTCTGTGCCGAGCTACTT---CGACTGCAGTCTCCAGCGCTGCGCAGGGGAGATCA 1794
Oy      1528  GAGAGGCCCTTCACCTGTGAGAAAGACGGCACCCTACTACCTGCTATAGGATPACTGACGTGG 1587
Db      1795  GGGGGGCCCCCTTGCCCTTGCAGAAACAGCGCTGCGCTTACCTTAGCGGATCATCAAGCTGG 1854
Oy      1588  GGCCCTGAGAGTGTGAG-----AAGAGCGCAGGGGCTCTACACCCAGCTTACCAATTCCTG 1641
Db      1855  GGTAGACGGCTGCGGGCGGCTCTCCAAAGCCGGGGGCTACACCGCGGTGCAAACTATGTG 1914
Oy      1642  AATTGATCAAGCCACCAT 1661
Db      1915  GACTGGATCAAGCAGCCGAT 1934

RESULT 2
US-08-448-937A-14
? Sequence 14, Application US/08448937A
? Patent No. 5677164
?
? GENERAL INFORMATION:
?
? APPLICANT: Takeshi SHIMOMURA et al.
?
? TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
?
? NUMBER OF SEQUENCES: 14
?
? CORRESPONDENCE ADDRESS:
?
? ADDRESSEE: Wenderoth, Lind & Ponack
?
? STREET: 805 Fifteenth Street, N.W., #700
?
? CITY: Washington
?
? STATE: D.C.
?
? COUNTRY: U.S.A.
?
? ZIP: 20005
?
? COMPUTER READABLE FORM:
?
? MEDIUM TYPE: Diskette, 5.25 inch,
?
? MEDIUM TYPE: 500 Kb Storage
?
? COMPUTER: IBM Compatible
?
? OPERATING SYSTEM: MS-DOS
?
? SOFTWARE: Wordperfect
?
? CURRENT APPLICATION DATA:
?
? APPLICATION NUMBER: US/08/448,937A
?
? FILING DATE: May 24, 1995
?
? CLASSIFICATION: 435
?
? PRIOR APPLICATION DATA:
?
? APPLICATION NUMBER: 08/148,910
?
? FILING DATE: No. 5677164ember 5, 1993
?
? ATTORNEY/AGENT INFORMATION:
?
? NAME: Warren M. Cheek, Jr.
?
? REGISTRATION NUMBER: 33,367
?
? REFERENCE/DOCKET NUMBER:
?
? TELECOMMUNICATION INFORMATION:
?
? TELEPHONE: 202-371-8850
?
? TELEFAX: 202-371-8856
?
? TELEX:
?
? INFORMATION FOR SEQ ID NO: 14:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 2033 base pairs
?
? TYPE: nucleic acid
?
? STRANDEDNESS: double
?
? TOPOLOGY: linear
?
? MOLECULE TYPE: cDNA
?
? ORIGINAL SOURCE:
?
? ORGANISM: human
?
? IMMEDIATE SOURCE:
?
? LIBRARY: Pre-made Lambda phage library,
?
? LIBRARY: human liver(49, male) cDNA library (Stratagene)
?
? US-08-448-937A-14
?
? Query Match 6.8%: Score 114.4; DB 1: Length 2033;
? Best Local Similarity 49.6%; Pred. No. 3.5e-24;
? Matches 605; Conservative 0; Mismatches 561; Indels 54; Gaps 10;
Oy      469  AACCCCTCCCAAGATGGGGCTACTGCTGCCGATTAAGCGGAGATCCAGTTGACCTGT 528
Db      742  AGCCCTCCCTCGAAGCGGGGACACCTGATTCGTGGCCACCGGACACCATCGTGTG 801

```

RESULT 2
 US-08-448-937A-14
 ; Sequence 14, Application US/08448937A
 ; Patent No. 5677164
 GENERAL INFORMATION:
 APPLICANT: Takeshi SHIMOMURA et al.
 TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wenderoth, Lind & Ponack
 STREET: 805 fifteenth street, N.W., #700
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 5.25 inch,
 MEDIUM TYPE: 500 Kb Storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/448,937A
 FILING DATE: May 24, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/148,910
 FILING DATE: No. 5677164ember 5, 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Warren M. Cheek, JF.
 REGISTRATION NUMBER: 33,367
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-8850
 TELEFAX: 202-371-8856
 TELEX:
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2033 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 ORGANISM: human
 IMMEDIATE SOURCE:
 LIBRARY: Pre-made Lambda phage library,
 LIBRARY: human liver(49, male) cDNA library (Stratagene)
 US-08-448-937A-14
 Query Match 6.8%: Score 114.4; DB 1: Length 2033;
 Best Local Similarity 49.6%; Pred. No. 3.5e-24;
 Matches 605; Conservative 0; Mismatches 561; Indels 54; Gaps 10.

```

OY 529 GCCTGTCGCCAGCAGTTCAAGGGAATTCCTGTAATAGTTTCATGA---CTGCTAT 585
    ||||| || ||||| ||||| ||||| ||||| ||||| |||||
DB 802 GCCGCGCCACAGAGGCTTACGAGGAGCGCTCTGCACATCGACCTGTGATGCGCTTC 861
OY 586 GTTGGCGATGGCTACTCTTACCGAGGGAATATAGACATGCATCAACACATGCGCTGC 645
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 862 TTGGGGAGAGCGCATGGGTACCGTGGGGTGGCCAGCAGCCTCAGGCTCGGGGCTCAGCTGC 921
OY 646 CTTTACTGGAACCTCCACCTCCTCTTGAGAGAGAAATTACAACTGTTATGASAGATGCT 705
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 922 CTGGCCCTGGAACCTCGATCTCTTACAGAGAGCTCGAGCTCGGTGGGCGCGCG 981
OY 706 GAAACCCATGGGATTGGGGAACACAAATTTCTGCAGAAACCCAGATGCGGAGAAAGCCC 765
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 982 GCCCTGTGGGGCTGGGGCCCGGATGCTGCTGCGGGAATCGGACAAATGACGAGAGGCC 1041
OY 766 TGGTCTTTATTAAGTTACCAATGACAGGTGAATGGGAATGCTGTGATGCTCAGCC 825
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1042 TGGTGTCTACGT---GGTGAAGAGACAGCGCGCTCCGAGGAGTACTGCCCTGAGAGCC 1098
OY 826 TGCTCAGCCAGGAGCGTTGCTTACCCAGAGGAAAGCCCACTGAGCATCAACCAAGCTT 885
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1099 TCGCAATCCCT-----CACCGAGTCCAACTGTCAACCGGATCTCTGGCAGCCCTG 1149
OY 886 CCGGGGTTTGACTCTCTGTGGAAGACTGAGATGACAGAGAGAAATCAAGAAATCTAT 945
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1150 CTTGAGCCAGCCCTCCCGGGGCGCAGCGCTGCGGAGGACAGAAAGAGGAGCGTTC 1209
OY 946 GGAAGCTTTTAAGACAGCGGCGGCAAGCACCATGAGCGCTCCAGTCCGCTG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1210 CTGGCGGCAGATGATGCGGGCGCTCCCTCCGCTGCGCCGCTCGACCCCTG----- 1262
OY 1006 CCTGTACACATCTCCATGCGCCCGAGGCGACTTCTGTGTGGGCGCTGATCCACCCCTGC 1065
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1263 --GCTGGCGCCCATCTCAATCGGAGCAGCTTCTGCGCGGAGCGCTGCTCCACACCTGC 1320
OY 1066 TGGGTGCTCACTGCTGCTGACACGACATGAATAAACAGACATCTAA-----AGGCTG 1119
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1321 TGGGTGGTGTGCGCGCCGACCTGCTTCCACAGCCCCCGAGGAGACGCTGCTCCGCTG 1380
OY 1120 GTGCTAAGGAGACGAGACCTGTGAAGAAAGAAATTTATGAGAGAGCTTTGAGGTGAG 1179
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1381 GTGCTGGGCGAGCATTTCTTCAACCGCAGCAGCATGTGACGAGCATTTGGGCACTGAG 1440
OY 1180 AAGATATTCAAGTACAGCCACTCAATGAAGAGATGATCCCAATGATATTGCA 1239
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1441 AAGTACATCCCGTACACCTGATCGGTGTTCAACCCAG---CGACCAACAGCTCGTC 1497
OY 1240 TTGCTCAAGTTAAAGCATGATGATGCTGCTCTAGAAATCCAAATACCTGAAGACT 1259
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1498 CTGATTCGGGCTGAAGAAAGGAGGAGCGGTGCAACAGCTGCGATGCGAGGCC 1557
OY 1300 GTTGTCTTGGCTGATG-----GATCTTTCCTCTGAGAGTGAATCTCTCTGCTC 1353
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1558 ATMTGCTGCGCCGAGCCCGGACGACACTTCCCGCAGGACACAAGTGCAGATGCGGGC 1617
OY 1354 TGGGG-----TGTACAGAAACAGAGAAAGGATCCCGGACGCTCTGATGCCAAATC 1407
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1618 TGGGGCCACTTGGATGAAGCGTGAAGCGCTACAGCTCCCTGGGAGAGGCCCTGCTG 1677
OY 1408 AAGGTGATTGCAACACTTTGTGCAACTCCCGCAACTCTATGACCACTGATTTGATGAC 1467
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1678 CCCCTGGTGGCGGACCAAGTGCAGACCCCTGAGGCTTACGAGCCGAGATCAGCCCC 1737
OY 1468 AGTAAATCTGTGCGAGAAATCTTCAGAAACCTGGGCAACACACTGCGAGGAGTACTCT 1527
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1738 AACATGCTGTGTGCGGCTACTT---CGATGCAAGTGCAGAGCTGACAGGAGGAGCTCA 1794
OY 1528 GGAGGCCCCCTGACCTGTGAGAAGAGCGCACACTACTAGCTATGAGTACTGAGCTGG 1597
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1795 GGGGGGCCCCGCTGCGTGCAGAAAGCGGCTGCTTACCTTACGAGCATCATCAGCTGG 1854

```

```

OY 1588 GGCCTGAGTGTGAG-----AAGAGCGCAGGGGCTTACACCAAGTTACCAAAATTCCTG 1641
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1855 GGTGACGCGCTCGGGGCGGCTCCACAAAGCCGGGGGTCTTACACCGCGTGGCAACTATGTG 1914
OY 1642 AATTGATCAAGCCACCAT 1661
    ||||| ||||| ||||| ||||| ||||| |||||
DB 1915 GACTGATCAACGACCGGAT 1934

```

RESULT 3

```

US-08-148-910-3
? Sequence 3, Application US/08148910
? Patent No. 5466593
? GENERAL INFORMATION:
? APPLICANT: Takeshi SHIMOMURA et al.
? TITLE OF INVENTION: No. 5466593el Protein and Gene Encoding Said Protein
? NUMBER OF SEQUENCES: 14
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Wenderoth, Lind & Ponack
? STREET: 805 Fifteenth Street, N.W., #700
? CITY: Washington
? STATE: D.C.
? COUNTRY: U.S.A.
? ZIP: 20005
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 5.25 inch,
? MEDIUM TYPE: 500 Kb Storage
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: MS-DOS
? SOFTWARE: Wordperfect
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/148,910
? FILING DATE: No. 5466593ember 5, 1993
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Marten M. Cheek, Jr.
? REGISTRATION NUMBER: 33,367
? REFERENCE/DOCKET NUMBER:
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-371-8850
? TELEFAX: 202-371-8856
? TELEX:
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 970 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? ORIGINAL SOURCE:
? ORGANISM: human
? IMMEDIATE SOURCE:
? LIBRARY: Pre-made Lambda phage Library, human liver (49, male)
? LIBRARY: cDNA Library (Stratagene)
US-08-148-910-3

Query Match 5.5%: Score 93; DB 1; Length 970;
Best Local Similarity 52.7%: Pred. No. 6,1e-18;
Matches 357; Conservative 0; Mismatches 290; Indels 30; Gaps 6;
OY 1009 CTACCATCTTCATATCCCGCAGGCACTTCTGTGTGGGCGCTGATCCACCCCTGCTGG 1068
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 201 CTGGCGGCATCTACATATGGGAGACGTTCTGCGCGGAGGAGCTGTGTCACACCTGCTGG 260
OY 1069 GTGCTACTGTGCGCCACTGCACCGACATATAAAGCAGACA-----TCTAAAGTGTGG 1122
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 261 GTGATGTGTGGCGCGCCACTGCTTCTCCACAGCCGCCAGGAGACGCGTTCGCTGAGTG 320
OY 1123 CTAGGAGCAGGACCTGAAGAAAGAAATTTCAATGACAGAGCTTTAGGCTGACAGAG 1182
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 321 CTGGGCGACAGCTTTCTTCAACCGACGACGAGCTGACGACACCTTCGGCATCGAAG 380
QY 1183 ATATTCAAGTACAGCTTCAATGAAAGATGAGATTCCCAATGATATTCATTG 1242
Db 381 TACATCCGCTACACCTGCTGCTGCTTCAACCCCA---CGACACGACCTGCTGCG 437
QY 1243 CTCAGATTAAAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1302
Db 438 ATCCGCTGAAGAAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
QY 1303 TCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356
Db 498 TCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 557
QY 1357 GG-----TGTACAGAAACAGAAAGGCTCCGCTGCTGCTGCTGCTGCTGCTG 1410
Db 558 GGCACCTGTGATGAAACGTGACGCGCTACTCCACTCCCTGCGGAGGCTTGTGCTCC 617
QY 1411 CTGATTGCGACACTTTGTGCACTCCGCGCACTGCTGCTGCTGCTGCTGCTGCTG 1470
Db 618 CTGCTGCGCGACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
QY 1471 ATGATCTGTGAGAAATCTTCAAGAACTGCGCAAGACACTGCTGCTGCTGCTG 1530
Db 678 ATGCTCTGTGCGGCTACTT---CGACTGCAAGTCCGACGCTGCTGCGGAGCTGAG 734
QY 1531 GGGCCCTGACCTGTGAGAGGAGGACCTTACTGCTGCTGCTGCTGCTGCTGCTG 1590
Db 735 GGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 794
QY 1591 CTGAGCTGTAG-----AAGAGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1644
Db 795 GAGGCTCGGGGCGCTCCACAAGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
QY 1645 TGGATCAAGCCACCAT 1661
Db 855 TGGATCAAGCCGAGT 871

RESULT 4
US-08-448-937A-3
; Sequence 3, Application US/08448937A
; Patent No. 5677164
; GENERAL INFORMATION:
; APPLICANT: Takeshi SHIMOMURA et al.
; TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch,
; MEDIUM TYPE: 500 Kb Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448, 937A
; FILING DATE: May 24, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/148, 910
; FILING DATE: No. 5677164member 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850

TELEFAX: 202-371-8856
; TELTEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; LIBRARY: Pre-made Lambda phage library, human liver (49, male)
; LIBRARY: cDNA library (Stratagene)
US-08-448-937A-3

Query Match 5.5%; Score 93; DB 1; Length 970;
Best Local Similarity 52.7%; Pred. No. 6.1e-18;
Matches 357; Conservative 0; Mismatches 290; Indels 30; Gaps 6;

QY 1009 CTGACCATCTCCATGCCCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
Db 201 CTGGCGCCCATCTACATCGGGAGACGCTTCTGCGCGGGAGCGCTGCTGCTGCTGCTG 260
QY 1069 GTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1122
Db 261 GTGGTGTGAGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 320
QY 1123 CTAGGGGACACGACCTGTAAGAAAGAAATTTCTATGACGACGCTTATGAGTGCAGAG 1182
Db 321 CTGGCGCCGACACTTCTTCAACGCGACGAGGAGCTGACGACGACCTTGGCATGGAAG 380
QY 1183 ATATTCAAGTACAGCTTCAATGAAAGATGAGATTCCCAATGATATTCATTG 1242
Db 381 TACATCCGCTACACCTGCTGCTGCTTCAACCCCA---CGACACGACCTGCTGCG 437
QY 1243 CTCAGATTAAAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1302
Db 438 ATCCGCTGAAGAAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
QY 1303 TCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356
Db 498 TCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 557
QY 1357 GG-----TGTACAGAAACAGAAAGGCTCCGCTGCTGCTGCTGCTGCTGCTG 1410
Db 558 GGCACCTGTGATGAAACGTGACGCGCTACTCCACTCCCTGCGGAGGCTTGTGCTCC 617
QY 1411 CTGATTGCGACACTTTGTGCACTCCGCGCACTGCTGCTGCTGCTGCTGCTGCTG 1470
Db 618 CTGCTGCGCGACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
QY 1471 ATGATCTGTGAGAAATCTTCAAGAACTGCGCAAGACACTGCTGCTGCTGCTG 1530
Db 678 ATGCTCTGTGCGGCTACTT---CGACTGCAAGTCCGACGCTGCTGCGGAGCTGAG 734
QY 1531 GGGCCCTGACCTGTGAGAGGAGGACCTTACTGCTGCTGCTGCTGCTGCTGCTG 1590
Db 735 GGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 794
QY 1591 CTGAGCTGTAG-----AAGAGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1644
Db 795 GAGGCTCGGGGCGCTCCACAAGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
QY 1645 TGGATCAAGCCACCAT 1661
Db 855 TGGATCAAGCCGAGT 871

RESULT 5
US-08-148-910-2
; Sequence 2, Application US/08148910
; Patent No. 546593

```

: GENERAL INFORMATION:
: APPLICANT: Takeshi SHIMOMURA et al.
: TITLE OF INVENTION: No. 546593el Protein and Gene Encoding Said Protein
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wenderoth, Lind & Ponack
: STREET: 805 Fifteenth Street, N.W., #700
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5.25 inch,
: MEDIUM TYPE: 500 Kb Storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/148,910
: FILING DATE: No. 546593ember 5, 1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Warren M. Cheek, Jr.
: REGISTRATION NUMBER: 33,367
: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-8850
: TELEFAX: 202-371-8856
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 329 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: human
: IMMEDIATE SOURCE:
: LIBRARY: Quick-cloneTM human liver cDNA (Clontech)
: US-08-148-910-2

Query Match          4.3%; Score 72.2; DB 1; Length 329;
Best Local Similarity 59.3%; Pred. No. 5,7e-12;
Matches 163; Conservative 0; Mismatches 103; Indels 9; Gaps 2;

QY 1381 TCCCGCCAGCTCCTGGATGCCAAGTCAAGCTGATTGCCAACCTTTGTGCACCTCCCGC 1440
DB 52 TCCAGCTCCCTCGCGGAGGCGCTGCTCCCTGTCGCCGACCAACAGTGCAGCGCCT 111
QY 1441 CAACCTATGACACATATGATGACAGTATGATCTGTGCAGGAAATCTTCAGAACT 1500
DB 112 GAGGTCTACGGCGCCGACATCAAGCCCAACATGCTGTGTGCGGCTACTT---CGACTGC 168
QY 1501 GGGCAAGACACTGCCAGGCTGACTCTGTGAGGCCCTCGACTCTGTGAGAAGACGGCACC 1560
DB 169 AAGTCGACGCTCCGACGAGGAGCTGAGGGGAGCCCTGCTGCGCAGAGAAGACGGCCTG 228
QY 1561 TACTACGCTATGAGATAGTGAAGCTGGGGCTGTGAGAGTGTGAG-----AAGAGCCAGAG 1614
DB 229 GCTTACCTCTACGGCATCATCACTGAGGTGACGGCTGGGGGCGCTCCACCAAGCGGGG 288
QY 1615 GTCTACACCCCAAGTTACCAAAATTCCTGAATTGGAT 1649
DB 289 GTCTACACCCGCGTGCCAACTATGTGACTGGAT 323

RESULT 6
US-08-148-910-13
; Sequence 13, Application US/08148910
```

```

: Patent No. 546593
: GENERAL INFORMATION:
: APPLICANT: Takeshi SHIMOMURA et al.
: TITLE OF INVENTION: No. 546593el Protein and Gene Encoding Said Protein
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wenderoth, Lind & Ponack
: STREET: 805 Fifteenth Street, N.W., #700
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5.25 inch,
: MEDIUM TYPE: 500 Kb Storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/148,910
: FILING DATE: No. 546593ember 5, 1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Warren M. Cheek, Jr.
: REGISTRATION NUMBER: 33,367
: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-8850
: TELEFAX: 202-371-8856
: TELEX:
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 329 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: human
: IMMEDIATE SOURCE:
: LIBRARY: Quick-cloneTM human liver cDNA (Clontech)
: US-08-148-910-13

Query Match          4.3%; Score 72.2; DB 1; Length 329;
Best Local Similarity 59.3%; Pred. No. 5,7e-12;
Matches 163; Conservative 0; Mismatches 103; Indels 9; Gaps 2;

QY 1381 TCCCGCCAGCTCCTGGATGCCAAGTCAAGCTGATTGCCAACCTTTGTGCACCTCCCGC 1440
DB 52 TCCAGCTCCCTCGCGGAGGCGCTGCTCCCTGTCGCCGACCAACAGTGCAGCGCCT 111
QY 1441 CAACCTATGACACATATGATGACAGTATGATCTGTGCGAAGAAATCTTCAGAACT 1500
DB 112 GAGGTCTACGGCGCCGACATCAAGCCCAACATGCTGTGTGCGGCTACTT---CGACTGC 168
QY 1501 GGGCAAGACACTGCCAGGCTGACTCTGTGAGGCCCTCGACTCTGTGAGAAGACGGCACC 1560
DB 169 AAGTCGACGCTCCGACGAGGAGCTGAGGGGAGCCCTGCTGCGCAGAGAAGACGGCCTG 228
QY 1561 TACTACGCTATGAGATAGTGAAGCTGGGGCTGTGAGAGTGTGAG-----AAGAGCCAGAG 1614
DB 229 GCTTACCTCTACGGCATCATCACTGAGGTGACGGCTGGGGGCGCTCCACCAAGCGGGG 288
QY 1615 GTCTACACCCCAAGTTACCAAAATTCCTGAATTGGAT 1649
DB 289 GTCTACACCCGCGTGCCAACTATGTGACTGGAT 323

RESULT 7
US-08-148-937A-2
```

```
; Sequence 2, Application US/08448937A
; Patent No. 5677164
; GENERAL INFORMATION:
; APPLICANT: Takeshi SHIMOMURA et al.
; TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch,
; MEDIUM TYPE: 500 Kb Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,937A
; FILING DATE: May 24, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/148,910
; FILING DATE: No. 5677164elember 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 329 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; LIBRARY: Quick-cloneTM human liver cdna (Clontech)
; US-08-448-937A-2

Query Match          4.3%; Score 72.2; DB 1; Length 329;
Best Local Similarity 59.3%; Pred. No. 5.7e-12;
Matches 163; Conservative 0; Mismatches 103; Indels 9; Gaps 2;
```

```
US-08-448-937A-13
; Sequence 13, Application US/08448937A
; Patent No. 5677164
; GENERAL INFORMATION:
; APPLICANT: Takeshi SHIMOMURA et al.
; TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch,
; MEDIUM TYPE: 500 Kb Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,937A
; FILING DATE: May 24, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/148,910
; FILING DATE: No. 5677164elember 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 329 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; LIBRARY: Quick-cloneTM human liver cdna (Clontech)
; US-08-448-937A-13

Query Match          4.3%; Score 72.2; DB 1; Length 329;
Best Local Similarity 59.3%; Pred. No. 5.7e-12;
Matches 163; Conservative 0; Mismatches 103; Indels 9; Gaps 2;
```



```
RESULT 9
US-08-427-640-1
; Sequence 1, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,640
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/689,410
; FILING DATE: 22 APRIL 1991
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1065 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: DNA
US-08-427-640-1

Query Match          4.3%: Score 72.2; DB 1; Length 1065;
Best Local Similarity 50.5%: Pred. No. 1,1e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

OY 1028 AGGGCCACTTCTGTGCTGGGGCGCTGATCCACCCCTGCTGGTGCATGCTGCCACT 1087
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 332 AGCGTTCCGTGTGGGGGCATCATCATCTCTGCTGATTCCTGTGCGGCCACT 451
OY 1088 GCACCGA-----CATAAAACCAACATCTAAAGTGTGCTGAGGGACCGACCTGA 1141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 452 GCTTCCAGGAGAGCTTTCGCCGCCACACCTGAGCGTGTCTTGGGACACATACCGGG 511
OY 1142 AGAAGAAGATTTCATGACGAGACCTTTAGGGTGCAGAAGATATTCAAGTACAGCCACT 1201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 512 TGGTCCCTGGCGAGGAGAGAGAAATTGAAGTCGAAAAATTCATTCATTAAGGAT 571
OY 1202 ACATTAAGAAAGATGAGATTCGCCACAAATGATATGATTCCTCAAGTTAAAGCCACTG 1261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 572 TCGAT-----GATGACACTTACGACATGACATTTGCTGCTGCGAGCTGAATCGGATT 625
OY 1262 ATGCTACTGTGCTAGAAATCAAAATACGTGAAGACTGTGTGCTTGCCTAGTGGTCTCT 1321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 626 CGTCCCGCTGTGCCAGSAGACACCGTGTGCGCACTGTGTCCCTCCCGCGGGAGCC 685
OY 1322 TTCCT-----CTGGAGTGTAGTSCCAATCTCTGGCTGGGTGGTTACAGAAA 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 686 TGCAGCTGCCGAGCGAGAGAGTGTGAGCTTTCGGCTACGCAAGCATGAGGCTTTGT 745
OY 1370 CAGGAAAAGGTCGCCGACCTCTCGATGCAAGTCAAGCTGAATTCGCAACACTTTGT 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 746 CTCCTTCTATTGCGAGCGGCTGAAGAGCTCATGTGACAGCTGATACCATTCAGCGCT 805
OY 1430 GCAACTCCCGCCCAACTATGACCATGATGATGATGATGATGATGATGATGATGATGATG 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 806 GCACATACAACTTTACTTAACAGAACAGTCAACGACACATGCTGTGTGCTGGAGCA 865
OY 1490 TTCGAACCTG-----GCAAGACACTGCCAGGGTGAATCTGTGAGGCC 1534
```

```
DB 866 CTGCGAGCGCGCGGCCCGCCAGCAACTTGGCAGACCCCTGCGAGGCGCATTTGGGAGGCC 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1535 CCTGACCTGTGAGAAGAGCGGACCTACTAGCTATGAGTATGAGTATGAGTATGAGTATG 1594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926 CCTGTGTGTCTGAACCATGCGCGCATGACTTGTGGTGATCATCATGCTGAGGCGCTGG 985
OY 1595 AGGTGAGAAGAG-----CCAGGGCTTACACCAAGTATCCAAATTCCTGAAATGCA 1648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 986 GCTGTGACAGAGAGATGTCGCCGGTGTGTACACAGGTATCAACTACTGACTGAG 1045
OY 1649 TCAAGCCACCAT 1651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1046 TTCGTGACAACT 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-08-427-640-5
; Sequence 5, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,640
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/689,410
; FILING DATE: 22 APRIL 1991
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1065 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-427-640-5

Query Match          4.3%: Score 72.2; DB 1; Length 1065;
Best Local Similarity 50.5%: Pred. No. 1,1e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

OY 1028 AGGGCCACTTCTGTGCTGGGGCGCTGATCCACCCCTGCTGGTGCATGCTGCCACT 1087
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 392 AGGGTCTCTGTGGGGGCATCATCATCTCTGCTGATTCCTGTGCGGCCACT 451
OY 1088 GCACCGA-----CATAAAACCAACATCTAAAGTGTGCTGAGGGACCGACCTGA 1141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 452 GCTTCCAGAGAGGTTTCCGCCACCACCTGACGATGATCTTGGCGAACAATACCGGG 511
OY 1142 AGAAGAAGATTTCATGACGAGAGCTTTAGGGTGCAGAAGATATTCATTAAGTACAGCCACT 1201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 512 TGGTCCCTGGGAGAGAGACAGAAATTTGAAGTGAAGAAATATGATTTGCTATAGGAAT 571
OY 1202 ACAATGAAGAAGATGATTCGCCACATGATATTTGCAATGCTCAAGTTAAAGCCAGTGG 1261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 572 TCGAT-----GATGACACTTACGACAAATGACATTTGCCCTGCTCACTGAATCGGATT 625
OY 1262 ATGCTACTGTGCTAGAAATCAAAATACGTGAAGACTGTGTGCTTCCCTGATGGTCTCT 1321
```

Db 626 CATTCCGCGTGTCCCAAGAGAGCAGGGTGATGATCCGCACTGTGTGCTTTCCCCCGGGGAGCC 685

Oy 1322 TTTCCT-----CTGGAGTGTAGTGCCACATTTCTTGGCTGGGGGTGTTACAGAAA 1369

Db 686 TGGAGCTGCCGAGCTGCAGCAGAGTGTGACCTCTCCGGCTACGGCAACATAGAGGCCCTGT 745

Oy 1370 CAGAGAAAGGATCCCGCCAGCTCTCTGGAGGCCAAAGTCAACCTCATTTGCCACACTTTGT 1429

Db 746 CTCTCTTTATTTCCGAGCGCGCTGAAGAGGCTCATGTGACTTACCCATTCAGCGCT 805

Oy 1430 GCACATCCCGCCCACTCATGACCACTGATTGATGACAGATGATCTGTGGCAGGAATC 1489

Db 806 GCACATCCCAACATTTACTTTACAGACAGTACCCGCAACATCTCTGTGGCTGGAGACA 865

Oy 1490 TTGAGAAACTGG-----GCAGAGACACTGCCAAGGTGACTCTGCAGGCC 1534

Db 866 CTCGGAGCGGGGCCCCAGGACAACTTGACAGACGCTCGCAAGGCGATTCCGGAGGCC 925

Oy 1535 CCGTGACCTGTGAGAAAGGACGCCACTACGTATGGGATGTAGTGTGGGGCTCGG 1594

Db 926 CCGCTGGTGTGTGCAGAGATGCCCCGACTGACTTTGGTGGCATATAGTGTGGGGCTTGG 985

Oy 1595 AGTGTGAGAAAGG-----CCAGGGGCTACACCCAAAGTTACCAAAATTCCTAAATTTGA 1648

Db 986 GGTGTGACAGAGAGATGTCCCGGGGTGTGACACCAAGATTTACCAACTCACTAGACTGA 1045

Oy 1649 TCACAGCCACAT 1661

Db 1046 TTCTGTGACACAT 1058

```

1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
190
```

Db	395	ACGGATTCTGTGCGGGGGGACATACATCATACAGTCTCTCTGTGATCTCTCTCTCCGCCACT	454
Oy	1088	GCACCGA-----CATAAAAACGACATCTAAAGGTGCGTAGGGGACAGACACTTA	1141
Db	455	GGTTCGAGGAGAGTTCCTCCGCCACCACCTACGCTGATCTTTGGGAGAAACATACCGG	514
Oy	1142	AGAAAGACATTTTATAGCAGAGCTTTAGGCTGCAGAAGATTTCAAGTACAGCCAT	1201
Db	515	TGTGCTCCGTGGCAGAGGAGAGAAATTTGAAGTCGAAATATCATATCTCCATAGAGAT	574
Oy	1202	ACATATGAAGAAGATGAGATTTCCCAATGATATTTGCAATTTGCTCAAGTTAAAGCCAGTGG	1261
Db	575	TGATAT-----GATGACACTTAGACAAATGATATTTGGCTGGCAGTGCATAAATCGGATT	628
Oy	1262	ATGGTCACTGTGCTTACAAATCCAAATACGTGAAGACGTGTGTCTTTGCTGATGGGTCT	1321
Db	629	CGTCCCGCTGTGCGCAGAGAGAGAGTGCTGCGACTGTGTGTCCTCCCGCCCGGCGACC	688
Oy	1322	TTCCCT-----CTGGAGTGAATGGCACATCTGTGGCTGGGGCTTACAGAAA	1365
Db	689	TGCAGTCTCCGAGCTGACAGGAGTGTGACTCTCTCCGCTACGGCAGCATAGAGCCCTTGT	748
Oy	1370	CAGGAAAGGGTCCCGCCAGCTCCTGTGATGCCAAAGTCAAGCTGATTTGCCAACCTTTGT	1429
Db	749	CTCTCTTCTATTTGACAGCGGCTGAAAGAGGCTCATGTACACTGTATCCATTCAGGCCCT	808
Oy	1430	GTAACTCCGCCCAACATCATACACATGATTAATGATGACATATGATCTGTGCAGAGAAATC	1489
Db	809	GCATCTACAACTTTACTTATACGAAAGAGTACCGSACACATGCTGTGTCTGTGGAGACA	868
Oy	1490	TTGAGAACTGG-----GCAAGACACCTGCGAGGGTCACTCTGTGAGGCC	1534
Db	869	CTCGAGAGGCGGGGGCCCCAGGCAAACTTCCACGAGAGGCTCCACGAGGCAATTCGGGAGCC	928
Oy	1535	CCCTGACTGTGGAAGAGACGCGACCTACTACTCTTATGGGATGTGAGCTGGGGCCGTGG	1594
Db	929	CCCTGTGTGTCTGAACATGAGCGCATACTTTGTGGGCAATCATAGCTGGGGCCGTGG	988
Oy	1595	ACTGTGAGAGAG-----CCAGGGGTCAACACCAGTTCACCAATTTCTGATTAATGA	1648
Db	989	GCTGTGAGACGAAGAGATGTCCGGGTGTATACACCAAGTTACCAATACCTACATCGGA	1048
Oy	1649	TCAAAAGCCACCAT 1661	
Db	1049	TTTCTGACACAT 1061	

```

RESULT 12
US-09-553-498-9
; Sequence 9, Application US/09553498
; Patent No. 6309861
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorothee
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwarz, Elisabeth
; TITLE OF INVENTION: Process for the production of naturally folded and secreted proteins
; FILE REFERENCE: Case 20379
; CURRENT APPLICATION NUMBER: US/09/553,498
; CURRENT FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: EP9107412.1
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 9
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: E. coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1137)
; US-09-553-498-9

```


GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 18:44:22 ; Search time 127.5 Seconds

(without alignments)
8816.543 Million cell updates/sec

Title: US-09-912-559-2

Perfect score: 1683
Sequence: 1 atgtttccagcagatgcttga.....aaagtgaaagtgcttctaa 1683

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 478924 seqs, 33395956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*\n2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*\n3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*\n4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*\n5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*\n6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:*\n7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*\n8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*\n9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*\n10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*\n11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*\n12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*\n14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683	100.0	1683	10	US-09-912-559-2
2	1679.8	99.8	1683	10	US-09-912-559-1
3	1679.8	99.8	3008	10	US-09-880-107-1668
C 4	120.2	7.1	451	10	US-09-864-761-11164
5	117	7.0	117	10	US-09-864-761-27791
6	114.4	6.8	2036	10	US-09-954-456-552
7	114.4	6.8	2036	10	US-09-880-107-1612
8	72.8	4.3	614	10	US-09-879-792-33
9	72.2	4.3	1609	10	US-09-969-271-6
10	72.2	4.3	2519	10	US-09-969-271-5
11	72.2	4.3	2641	9	US-09-974-298-144
12	66.4	3.9	1614	10	US-09-888-615-45
13	66.4	3.9	1748	10	US-09-879-792-11
14	66.4	3.9	2412	9	US-09-978-295A-68
15	66.4	3.9	2412	9	US-09-978-697-68
16	66.4	3.9	2412	9	US-09-978-192A-68
17	66.4	3.9	2412	9	US-09-999-832A-68
18	66.4	3.9	2412	9	US-09-978-189-68
19	66.4	3.9	2412	9	US-10-174-590-63

20	66.4	3.9	2412	9	US-10-176-758-63	Sequence 63, Appl
21	66.4	3.9	2412	9	US-10-175-737-63	Sequence 63, Appl
22	66.4	3.9	2412	9	US-10-173-706-63	Sequence 63, Appl
23	66.4	3.9	2412	9	US-10-175-738-63	Sequence 63, Appl
24	66.4	3.9	2412	9	US-10-175-752-63	Sequence 63, Appl
25	66.4	3.9	2412	9	US-10-176-482-63	Sequence 63, Appl
26	66.4	3.9	2412	9	US-10-176-757-63	Sequence 63, Appl
27	66.4	3.9	2412	9	US-10-176-913-63	Sequence 63, Appl
28	66.4	3.9	2412	9	US-10-180-552-63	Sequence 63, Appl
29	66.4	3.9	2412	9	US-10-180-557-63	Sequence 63, Appl
30	66.4	3.9	2412	9	US-10-173-700-63	Sequence 63, Appl
31	66.4	3.9	2412	9	US-10-174-572-63	Sequence 63, Appl
32	66.4	3.9	2412	9	US-10-174-579-63	Sequence 63, Appl
33	66.4	3.9	2412	9	US-10-174-582-63	Sequence 63, Appl
34	66.4	3.9	2412	9	US-10-174-588-63	Sequence 63, Appl
35	66.4	3.9	2412	9	US-10-175-739-63	Sequence 63, Appl
36	66.4	3.9	2412	9	US-10-175-740-63	Sequence 63, Appl
37	66.4	3.9	2412	9	US-10-175-743-63	Sequence 63, Appl
38	66.4	3.9	2412	9	US-10-176-488-63	Sequence 63, Appl
39	66.4	3.9	2412	9	US-10-176-492-63	Sequence 63, Appl
40	66.4	3.9	2412	9	US-10-176-747-63	Sequence 63, Appl
41	66.4	3.9	2412	9	US-10-176-750-63	Sequence 63, Appl
42	66.4	3.9	2412	9	US-10-176-985-63	Sequence 63, Appl
43	66.4	3.9	2412	9	US-10-176-987-63	Sequence 63, Appl
44	66.4	3.9	2412	9	US-10-176-991-63	Sequence 63, Appl
45	66.4	3.9	2412	9	US-10-176-992-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-912-559-2
; Sequence 2, Application US/09912559
; Patent No. US20020142316A1
; GENERAL INFORMATION:
; APPLICANT: ROEMER, JUERGEN
; APPLICANT: STOEHR, HANS-ARNOLD
; APPLICANT: FEUSSNER, ANNETTE
; APPLICANT: LANG, WIEGAND
; APPLICANT: WEIMER, THOMAS
; APPLICANT: BECKER, MARGRET
; APPLICANT: NERLICH, CLAUDIA
; APPLICANT: MUTH-NAUMANN, GUDRUN
; TITLE OF INVENTION: MUTANTS OF THE FACTOR VII-ACTIVATING PROTEASE AND
; TITLE OF INVENTION: DETECTION METHODS USING SPECIFIC ANTIBODIES
; FILE REFERENCE: 06478.1457
; CURRENT APPLICATION NUMBER: US/09/912.559
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: DE 100 36 641.4
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: DE 100 50 040.4
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: DE 100 52 319.6
; PRIOR FILING DATE: 2000-10-21
; PRIOR APPLICATION NUMBER: DE 101 18 706.8
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-912-559-2

Query Match 100.0%; Score 1683; DB 10; Length 1683;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTCCAGGATGCTGATCTCATCTTCTGCTGTATATGAGCTCTGCGGAAGACA 60
DB 1 ATGTTTCCAGGATGCTGATCTCATCTTCTGCTGTATATGAGCTCTGCGGAAGACA 60

[illegible]

Dd	1141	AAGAAAGAAGAAATTTTCATGACACAGAGCTTVAAGGTCCAGAAAGATTTCAGAGTACACCAC	1200
Oy	1201	TACATATGAAGAAGATGAGATATGCCCAACAATGATATGGCATTGCTCCTCAAGTTAAGGCAGTG	1260
Dd	1201	TACATATGAAGAAGATGAGATATGCCCAACAATGATATGGCATTGCTCCTCAAGTTAAGGCAGTG	1260
Oy	1261	GATGTGTACTGTGTCTGTGAAATCCAATTAAGTGAAGACTGTGTGTGCTTGCCTATGGGTCC	1320
Dd	1261	GATGTGTACTGTGTCTGTGAAATCCAATTAAGTGAAGACTGTGTGTGCTTGCCTATGGGTCC	1320
Oy	1321	TTTTCCCTGTGGAGTAGTGTGCACATCTGTGGCTGGGGGTATTACAGAAACAGAAAAGGG	1380
Dd	1321	TTTTCCCTGTGGAGTAGTGTGCACATCTGTGGCTGGGGGTATTACAGAAACAGAAAAGGG	1380
Oy	1381	TCCCGCACCTCTCTGGATGGCCAAAGTCAAGCTGATTTGCCAACACTTTTGTCAACTCCCG	1440
Dd	1381	TCCCGCACCTCTCTGGATGGCCAAAGTCAAGCTGATTTGCCAACACTTTTGTCAACTCCCG	1440
Oy	1441	CAACTGTATGACACAGATGATTTGATGACAGTATGATCTGTGAGGAAATCTTCAGAAACT	1500
Dd	1441	CAACTGTATGACACAGATGATTTGATGACAGTATGATCTGTGAGGAAATCTTCAGAAACT	1500
Oy	1501	GGGCAAGACACCTGTGCAGAGGTGATCTGTGAGAGCCCCCTGACTGTGCAAGAGAGCCGAC	1560
Dd	1501	GGGCAAGACACCTGTGCAGAGGTGATCTGTGAGAGCCCCCTGACTGTGCAAGAGAGCCGAC	1560
Oy	1561	TACTAGCTGTATGGATAGTAGAGCTGGGGCTGTGAGTGTGAGAAAGAGCCAGAGGTCTAC	1620
Dd	1561	TACTAGCTGTATGGATAGTAGAGCTGGGGCTGTGAGTGTGAGAAAGAGCCAGAGGTCTAC	1620
Oy	1621	ACCCAGTTACCAAATTTCTGATTTGATGATCAAAAGCCACCATCAAAATGAAAGTGGCTTC	1680
Dd	1621	ACCCAGTTACCAAATTTCTGATTTGATGATCAAAAGCCACCATCAAAATGAAAGTGGCTTC	1680
Oy	1681	TAA 1683	
Dd	1681	TAA 1683	
 RESULT 2 US-09-912-559-1			
Sequence 1, Application US/09912559			
Patent No. US20020142316A1			
GENERAL INFORMATION:			
APPLICANT: KOEMISCH, JUERGEN			
APPLICANT: STOEHR, HANS-ARNOLD			
APPLICANT: FEUSSNER, ANNETTE			
APPLICANT: LANG, WIEGAND			
APPLICANT: WEIMER, THOMAS			
APPLICANT: BECKER, MARGRET			
APPLICANT: NERLICH, CLAUDIA			
APPLICANT: MOTH-NORMANN, GUDRUN			
TITLE OF INVENTION: MUTANTS OF THE FACTOR VII-ACTIVATING PROTEASE AND			
FILE REFERENCE: 06478.1457			
CURRENT APPLICATION NUMBER: US/09/912.559			
CURRENT FILING DATE: 2001-07-26			
PRIOR APPLICATION NUMBER: DE 100 36 641.4			
PRIOR FILING DATE: 2000-07-26			
PRIOR APPLICATION NUMBER: DE 100 50 040.4			
PRIOR FILING DATE: 2000-10-10			
PRIOR APPLICATION NUMBER: DE 100 52 319.6			
PRIOR FILING DATE: 2000-10-21			
PRIOR APPLICATION NUMBER: DE 101 18 706.8			
PRIOR FILING DATE: 2001-04-12			
NUMBER OF SEQ ID NOS: 4			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 1			
LENGTH: 1683			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-912-559-1			

Query Match 99.8%; Score 1679, 8; DB 10; Length 1683;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGTTTGCCAGAGATGTCGATCTCCATGTTCTGTTAATGGCTGTGGGAAAGACA 60
Db 1 ATGTTTGCCAGAGATGTCGATCTCCATGTTCTGTTAATGGCTGTGGGAAAGACA 60
QY 61 GCCTGGGGTTCCTCCATGTTCTTTATTTGGAAGCCGAGCCGAGCTGACCCCTGAC 120
Db 61 GCCTGGGGTTCCTCCATGTTCTTTATTTGGAAGCCGAGCCGAGCTGACCCCTGAC 120
QY 121 CAGTATGATTACAGCTACGAGGATTAATACAGAGAACACCAAGTACACACTTACC 180
Db 121 CAGTATGATTACAGCTACGAGGATTAATACAGAGAACACCAAGTACACACTTACC 180
QY 181 CATGCTGAGAAATCCGACTGCTACTACAGAGACCAAGCTGATCCATGCGACGCCAAC 240
Db 181 CATGCTGAGAAATCCGACTGCTACTACAGAGACCAAGCTGATCCATGCGACGCCAAC 240
QY 241 CCCTGTGAACACGGTGGGAGTGCCTGCTGCATGGAGCACCCTTACATGCACTGCTTG 300
Db 241 CCCTGTGAACACGGTGGGAGTGCCTGCTGCATGGAGCACCCTTACATGCACTGCTTG 300
QY 301 GCTCCTTTCTGGGAATAGTGTCAAGAAAGTCAAAATACGTGCAAGAGAACACCATGT 360
Db 301 GCTCCTTTCTGGGAATAGTGTCAAGAAAGTCAAAATACGTGCAAGAGAACACCATGT 360
QY 361 GCGCGGGGCAATGTCATATTACCCAGAGTCCCTCCATACCGCTGCTGTCTGTAACAC 420
Db 361 GCGCGGGGCAATGTCATATTACCCAGAGTCCCTCCATACCGCTGCTGTCTGTAACAC 420
QY 421 CCTTACACAGTCCAGCTGCTCCCAAGTGTCTCTGATGAGAGCCAAACCCCTGACAG 480
Db 421 CCTTACACAGTCCAGCTGCTCCCAAGTGTCTCTGATGAGAGCCAAACCCCTGACAG 480
QY 481 AATGGGGTACTGCTGCCCGCATTAAGCCGAGATCCAGTTCACTGCTGCTGCCGAC 540
Db 481 AATGGGGTACTGCTGCCCGCATTAAGCCGAGATCCAGTTCACTGCTGCTGCCGAC 540
QY 541 CAGTTCAAGGGGAAATTTCTGAAATAGGTTGATGACTGCTGATGTTGGCGATGCTAC 600
Db 541 CAGTTCAAGGGGAAATTTCTGAAATAGGTTGATGACTGCTGATGTTGGCGATGCTAC 600
QY 601 TCTTACCGAGGAAATAGATAGACAGTCAACACAGCATGGTGCCTTTACTGGAACCTCC 660
Db 601 TCTTACCGAGGAAATAGATAGACAGTCAACACAGCATGGTGCCTTTACTGGAACCTCC 660
QY 661 CACTCTCTCTTGCAGAGAAATTACAACTGTTTATGAGAGATGCTGAACCCCATGGATT 720
Db 661 CACTCTCTCTTGCAGAGAAATTACAACTGTTTATGAGAGATGCTGAACCCCATGGATT 720
QY 721 GGGGAACACAAATTTCTGCAGAAACCCAGATGGGAGCAAAACCCCTGCTTATTTAAA 780
Db 721 GGGGAACACAAATTTCTGCAGAAACCCAGATGGGAGCAAAACCCCTGCTTATTTAAA 780
QY 781 GTTACCAATGACAAAGTGAATGGAATCTGTGATGCTCAGCCTGCTCAGCCAGGAC 840
Db 781 GTTACCAATGACAAAGTGAATGGAATCTGTGATGCTCAGCCTGCTCAGCCAGGAC 840
QY 841 GTTGGCTTCCAGAGAAAGCCCACTGAGCCATCAACAGCTTCCGGGGTTTACTGCC 900
Db 841 GTTGGCTTCCAGAGAAAGCCCACTGAGCCATCAACAGCTTCCGGGGTTTACTGCC 900
QY 901 TGTGGAAGACTGAGATAGCAGAGAGAAAGATCAAGAGAAATCTATGAGGCTTTAAGAC 960
Db 901 TGTGGAAGACTGAGATAGCAGAGAGAAAGATCAAGAGAAATCTATGAGGCTTTAAGAC 960
QY 961 ACGGGGGGCAAGCCCATGGAGGCTCCCTCCAGTCCCTGCTGCTGACCATCTCC 1020
Db 961 ACGGGGGGCAAGCCCATGGAGGCTCCCTCCAGTCCCTGCTGCTGACCATCTCC 1020
```

```
QY 1021 ATGCCCAAGGGCCACTTCTGTGTGGGGGCTGATCCACCCCTGCTGCTCACTGCT 1080
Db 1021 ATGCCCAAGGGCCACTTCTGTGTGGGGGCTGATCCACCCCTGCTGCTCACTGCT 1080
QY 1081 GCCCACTGACCGACATATAAAMACAGACATCTAAAGGTGGTGCAGAGAGACCTG 1140
Db 1081 GCCCACTGACCGACATATAAAMACAGACATCTAAAGGTGGTGCAGAGAGACCTG 1140
QY 1141 AAGAAAGAAATTTCTATGAGCAGAGCTTTAGGGTGCAGAGATATTCAAGTACGCCAC 1200
Db 1141 AAGAAAGAAATTTCTATGAGCAGAGCTTTAGGGTGCAGAGATATTCAAGTACGCCAC 1200
QY 1201 TACAAATGAAAGATGATGATCCCAATGATATGCAATGCTGCTCAAGTTAAAGCAGTG 1260
Db 1201 TACAAATGAAAGATGATGATCCCAATGATATGCAATGCTGCTCAAGTTAAAGCAGTG 1260
QY 1261 GATGGTCACTGTGTCTAGAAATCCAAATACGTGAAGACGTGTGCTTGCCTGATGGTCC 1320
Db 1261 GATGGTCACTGTGTCTAGAAATCCAAATACGTGAAGACGTGTGCTTGCCTGATGGTCC 1320
QY 1321 TTTCCCTGTGGAGTGAAGTGCACATCTCTGGCTGGGGTGTACAGAAACAGAAAGGG 1380
Db 1321 TTTCCCTGTGGAGTGAAGTGCACATCTCTGGCTGGGGTGTACAGAAACAGAAAGGG 1380
QY 1381 TCCCGCCAGCTCCGAGTGCACAAAGTCAAGCTGATGCAACCTTTGTGCAACCTCCGC 1440
Db 1381 TCCCGCCAGCTCCGAGTGCACAAAGTCAAGCTGATGCAACCTTTGTGCAACCTCCGC 1440
QY 1441 CAACTCTATGACCAATGATGATGATGACAGTATGATCTGTGTCAGAGAAATTTAGAAACT 1500
Db 1441 CAACTCTATGACCAATGATGATGATGATGACAGTATGATCTGTGTCAGAGAAATTTAGAAACT 1500
QY 1501 GGGCAGACACCTGCGAGGGTGAATCTGGAGCCCTGACCTGTGAGAGAGAGGACCC 1560
Db 1501 GGGCAGACACCTGCGAGGGTGAATCTGGAGCCCTGACCTGTGAGAGAGAGGACCC 1560
QY 1561 TACTACGTATGGGATAGTGAAGTGGGGCTGAGTGTGAGAGAGGCGAAGGGTCTAC 1620
Db 1561 TACTACGTATGGGATAGTGAAGTGGGGCTGAGTGTGAGAGAGGCGAAGGGTCTAC 1620
QY 1621 ACCCAAGTTACCAATTTCTGAAATGATGATCAAAAGCCACATCAAAAGTGAAGTGGCTTC 1680
Db 1621 ACCCAAGTTACCAATTTCTGAAATGATGATCAAAAGCCACATCAAAAGTGAAGTGGCTTC 1680
QY 1681 TAA 1683
Db 1681 TAA 1683

RESULT 3
US-09-880-107-1668
; Sequence 1668, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1668
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

OTHER INFORMATION: Genbank Accession No. US20020142981A1 D49742
US-09-880-107-1668

Query Match 99.8%; Score 1679.8; DB 10; Length 3008;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTTTGCCAGAGATGTCGATCTCCATGTTCTGCTGTATATGGCTCTGGTGGGAAACACA 60
DB 97 ATGTTTGCCAGAGATGTCGATCTCCATGTTCTGCTGTATATGGCTCTGGTGGGAAACACA 156
QY 61 GCGTGTGGGTTCTCCCTGATGTCCTTTATTTGAAAAGCTGGACCCAGACTGGACCCCTGAC 120
DB 157 GCGTGTGGGTTCTCCCTGATGTCCTTTATTTGAAAAGCTGGACCCAGACTGGACCCCTGAC 216
QY 121 CAGATGATTTACAGCTACAGAGATTTATATCAGAGAGAACACCAGTACACTTACC 180
DB 217 CAGATGATTTACAGCTACAGAGATTTATATCAGAGAGAACACCAGTACACTTACC 276
QY 181 CATGCTGAGATCTGACTGCTACTACTAGAGACCAGCTGATCCATGCCAGCCCAAC 240
DB 277 CATGCTGAGATCTGACTGCTACTACTAGAGACCAGCTGATCCATGCCAGCCCAAC 336
QY 241 CCCGTGTAAACAGGTGGGAGACTGCTGCTGATGGAGACACCTTCAATGAGCTGGCTG 300
DB 337 CCCGTGTAAACAGGTGGGAGACTGCTGCTGATGGAGACACCTTCAATGAGCTGGCTG 396
QY 301 GCTCTCTTCTGTGGGAATAGTGTAGAGAAAGTGC AAAATACGTGCAGAGACAAACCATGT 360
DB 397 GCTCTCTTCTGTGGGAATAGTGTAGAGAAAGTGC AAAATACGTGCAGAGACAAACCATGT 456
QY 361 GCGCGGGGCCAAATGTCTATTTACCCAGAGTCCTCCCTACTACCGCTGTCTGTAAACAC 420
DB 457 GCGCGGGGCCAAATGTCTATTTACCCAGAGTCCTCCCTACTACCGCTGTCTGTAAACAC 516
QY 421 CCTTACACAGATCCACAGCTGCTCCCAAGTGGTTCCTGTATGACAGGCCAAACCCCTGAC 480
DB 517 CCTTACACAGATCCACAGCTGCTCCCAAGTGGTTCCTGTATGACAGGCCAAACCCCTGAC 576
QY 481 AATGGGCTACCTGCTCCGCGGATTAAGCGGAGATCCAAAGTTCACCTGTCCTGCTCCGAC 540
DB 577 AATGGGCTACCTGCTCCGCGGATTAAGCGGAGATCCAAAGTTCACCTGTCCTGCTCCGAC 636
QY 541 CAGTTCAAGGGGAAATTTCTGTGAATAGTTCGTAGTACTGCTATGTTGGCGATGCTAC 600
DB 637 CAGTTCAAGGGGAAATTTCTGTGAATAGTTCGTAGTACTGCTATGTTGGCGATGCTAC 696
QY 601 TCTTACGAGGAGAAATGAATAGTACAGTCAACAGGATGCTGCTTACTGGAACCTCC 660
DB 697 TCTTACGAGGAGAAATGAATAGTACAGTCAACAGGATGCTGCTTACTGGAACCTCC 756
QY 661 CACCTCTCTTGCAGAGAAATTACACATGTTTATGGAGGATGCTGAAACCCATGGGATT 720
DB 757 CACCTCTCTTGCAGAGAAATTACACATGTTTATGGAGGATGCTGAAACCCATGGGATT 816
QY 721 GGGGAGACAAATTTCTGAGAAACCCAGATCGGACCAAAAGCTGGTGTATTATTA 780
DB 817 GGGGAGACAAATTTCTGAGAAACCCAGATCGGACCAAAAGCTGGTGTATTATTA 876
QY 781 GTTACCAATGACAAGGTGAATGGAATAGTGTATGCTACGCTGCTACGCCAGGAC 840
DB 877 GTTACCAATGACAAGGTGAATGGAATAGTGTATGCTACGCTGCTACGCCAGGAC 936
QY 841 GTTGCCTACCCAGAGGAAAGCCCACTGAGCATCAACCAAGTTCCGGGGTTTACCTCC 900
DB 937 GTTGCCTACCCAGAGGAAAGCCCACTGAGCATCAACCAAGTTCCGGGGTTTACCTCC 996
QY 901 TGTGAAAAGACTGATAGCAGAGAGAAAGATCAAGAATCTATGAGAGCTTTTAAGAC 960
DB 997 TGTGAAAAGACTGATAGCAGAGAGAAAGATCAAGAATCTATGAGAGCTTTTAAGAC 1056
QY 961 ACGGCGGAGAACCACTGAGGAGGCTCCCAAGTCTCGTGTGACCAATCTCC 1020

DB 1057 ACGGCGGAGAACCACTGAGGAGGCTCCCAAGTCTCGTGTGACCAATCTCC 1116
QY 1021 ATGCCCGAGGGCCACTTCTGTGTGGGGCGCTGATCCACCCCTGTGGTCTCACTCT 1080
DB 1117 ATGCCCGAGGGCCACTTCTGTGTGGGGCGCTGATCCACCCCTGTGGTCTCACTCT 1176
QY 1081 GCCCATGTACCCGATTAATAAACAGACATCTAAAGGTGTGTAGGGAGCAGGACCTG 1140
DB 1177 GCCCATGTACCCGATTAATAAACAGACATCTAAAGGTGTGTAGGGAGCAGGACCTG 1236
QY 1141 AAGAAAGAAATTTCTATGAGCAGAGCTTTAGGTGCAAGAATATTCAAGTACGCCAC 1200
DB 1237 AAGAAAGAAATTTCTATGAGCAGAGCTTTAGGTGAGAGATATTCAAGTACGCCAC 1296
QY 1201 TACATTAAGAGATGATGATTTCCCAATGATATTCATTCGCTCAAGTTAAAGCAGTGT 1260
DB 1297 TACATTAAGAGATGATGATTTCCCAATGATATTCATTCGCTCAAGTTAAAGCAGTGT 1356
QY 1261 GATGCTCACTGCTCTAGATCAATTAAGTGAAGACTGTGTGCTTGCCTGATGGCTCC 1320
DB 1357 GATGCTCACTGCTCTAGATCAATTAAGTGAAGACTGTGTGCTTGCCTGATGGCTCC 1416
QY 1321 TTTCCTCTGGAGTGAAGTGCACATCTCTGCTGGGTGTTACAGAAACAGGAAAGG 1380
DB 1417 TTTCCTCTGGAGTGAAGTGCACATCTCTGCTGGGTGTTACAGAAACAGGAAAGG 1476
QY 1381 TCCCGCCAGCTCTGTGATGCCAAAGTCAAGCTGATTCGCAACCTTTGTGCAACTCCGC 1440
DB 1477 TCCCGCCAGCTCTGTGATGCCAAAGTCAAGCTGATTCGCAACCTTTGTGCAACTCCGC 1536
QY 1441 CAACCTCTATGACCACTGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1537 CAACCTCTATGACCACTGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1596
QY 1501 GGGCAAGACACCTGCCAGGGTGACTGTGAGGCCCTCTGACTGTGAGAAAGACGGCAC 1560
DB 1597 GGGCAAGACACCTGCCAGGGTGACTGTGAGGCCCTCTGACTGTGAGAAAGACGGCAC 1656
QY 1561 TACTACGTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1657 TACTACGTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1716
QY 1621 ACCCAAGTTACCAATTTCTGATTTGGATCAAGCCACCATCAAAATGGAAGTGGCTTC 1680
DB 1717 ACCCAAGTTACCAATTTCTGATTTGGATCAAGCCACCATCAAAATGGAAGTGGCTTC 1776
QY 1681 TAA 1683
DB 1777 TAA 1779

RESULT 4
US-09-864-761-11164/c
; Sequence 11164, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US 09/1864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04


```

? Patent No: US20020115057A1
? GENERAL INFORMATION:
? APPLICANT: Young, Paul
? TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Candi
? TITLE OF INVENTION: Sets
? FILE REFERENCE: 689290-76
? CURRENT APPLICATION NUMBER: US/09/954,456
? CURRENT FILING DATE: 2001-09-18
? PRIOR APPLICATION NUMBER: US/60/233,617
? PRIOR FILING DATE: 2000-09-18
? PRIOR APPLICATION NUMBER: US/60/234,052
? PRIOR FILING DATE: 2000-09-20
? PRIOR APPLICATION NUMBER: US/60/234,923
? PRIOR FILING DATE: 2000-09-25
? PRIOR APPLICATION NUMBER: US/60/235,134
? PRIOR FILING DATE: 2000-09-25
? PRIOR APPLICATION NUMBER: US/60/235,637
? PRIOR FILING DATE: 2000-09-26
? PRIOR APPLICATION NUMBER: US/60/235,638
? PRIOR FILING DATE: 2000-09-26
? PRIOR APPLICATION NUMBER: US/60/235,711
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: US/60/235,720
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: US/60/235,840
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: US/60/235,863
? PRIOR FILING DATE: 2000-09-27
? NUMBER OF SEQ ID NOS: 2276
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 552
? LENGTH: 2036
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-954-456-552

```

Query Match	Local Similarity	6.83;	Score 114.4;	DB 10;	Length 2036;
Matches	605;	Conservative	0;	MisMatches 561;	Indels 54;
					Gaps 10.
QY	469	AACCCCTGCCAATATGAGGAGGCTACCTCTCTCCCGCATGAAGGAGATCCAAATGTCACCTGT	528		
Db	745	AGCCCTTGCTGAACGGGGGACACTGCACCTGATCGGGCCACGGGACCAACCGTGTGT	804		
QY	529	GCCGTGCCGACCAATTCACAGGGGAAATCTGTGAATAGTTCATATA---CTGCTAT	585		
Db	805	GCCGTGCCCAACAGGCTTCCTGGAACGGCTCTGCACATCAAGCTTATGAGCGCTGTCTC	864		
QY	586	GTTCGCGATGCTACTCTTACCGAGGAAATGATAGGACAGTCACACGATGATCCGTC	645		
Db	865	TTGGGGAMCGCATCGGTACCGCTGGCGCGAGACACCTCACGCTCGGGCTCGAGTCTC	924		
QY	646	CTTTACTGGAATCCACCTCTCTTGCGAGGAAATTACAACTTTTATGAGATGCT	705		
Db	925	CTGGCCTGGAACCTCCGATCTGCTCTTACCGAGGCTGCACTGAGATCTCGTGGCGCGCG	984		
QY	706	GAACCCATGAGATGGGGAACCAATTTCTGCAGAAACCAGATCCGAGAGAAAGGCC	765		
Db	985	GCCCTCTCGGGCTCGGGCCCCCATGTCTTACGCCGAATCCGACATAGCAGAGAGGCC	1044		
QY	766	TGCTGTTTATTAAGTTATTCATTCACAAGGAAATGGGAATACGTGATGTCTACGCC	825		
Db	1045	TGCTGCTACGT---GTGAAGGACAGCGCCTCTCTCTGGGAGTACGCGCCTGGAAGCC	1101		
QY	826	TGCTCAACCCACAGAGCTTGCCTACCCCAAGAGAAAGCCCACTGAGGCATCAACCACTT	885		
Db	1102	TGGGAATCCCT-----CACCAAGTCCCACTGTCAACGGATCTCTGTGACACCCCTG	1152		
QY	886	CCGGGGTTTGAATCTCTGTGAAGAAGTGTGATTCACAGAGGAAGTCAATGATGTAT	945		
Db	1153	CCGAGACGAGCTCCCGGGGGGCCCAAGGCTGTGGCAGAGAGGATCAAGAAGAGACCTTC	1212		
QY	946	GGAGGCTTTAAGACACGCGGGCAAGCAACCATATGACAGGCTCCCTCCAGTCTCTGCTG	1005		

Db 1213 CTGGGCGCAGCTATCAGGCGCTCTCTCCCTGCGCCGCGGTGCAACCCCTG----- 1265

Qy 1006 CCTGTGACCATCTCCATCGCCGACGGGCACTTGTGTGGGGGGGTGATCCACCCCTGC 1065

Db 1266 --GCTGGCCGCCATCTACATGTGGGAGACGTTCTTGCGCCGGGAGCTGTGTCCACACTGC 1323

Qy 1066 TGGGTCTCACTGTGCGCCACTGTGCACGGACATAAAAACAGCATCTAA-----AGGTG 1119

Db 1324 TGGTGTGTGTGGCGCGCCACTGCTCTCCACAGCCGCCACCGGAGACACTCTCCGTG 1383

Qy 1120 GTGTAGGGGACAGGACCTCTAACAAGAAATTTATTAAGACAGACTTTAGGGTGCAG 1179

Db 1384 GTGTGGGGCAGCACTTCTTCAACCGACAGACGAGCTGTGACAGACACTTTCGCACTGAG 1443

Qy 1180 AAGATATTCAAGTACAGCCACTACATGAAGAAAGATGAGATTCCCAATGAATATGCA 1239

Db 1444 AAGATACATCCGCTACACCCCTACTCGTGTTCACCCAG--CGACCCAGCACTCGTC 1500

Qy 1240 TTGCTCAAGTTAAACCGAGTGATGTGCACTGTGCTGATGAATCCAAATACGTAAGACT 1299

Db 1501 CTGATCGGGCTAABAACAAGGAGACGCTGTGCGACACGCTGCACTTCTGTGACGCC 1560

Qy 1300 GTGTGCTGCTGTGAG-----GCTCTTTCCCTGTGGAGTGAAGTCCACATCTTGTGC 1353

Db 1561 ATCTGCTCTCCGACCCCGGACGACCTTCTCCCGAGACACAAATGTCGAGATTGGGGTC 1620

Qy 1354 TGGGG-----TGTTACAGAAACAGGAAAGGGTCCCGCAGCTCTCGATGCGCAAGTC 1407

Db 1621 TGGGGCCACTTGTGATGAAGAAAGTAGGGGCTACTCAGCTTCCCTGTGGGAGGACCTGGTC 1680

Qy 1408 AAGCTGATTTGCCACACTTTGTGCACTCCCGCACTCTATGACCAATGATTGATGAC 1467

Db 1681 CCGCTGTGGCGACACACAGTAGTAGACCTGTGAGGCTTACGCGCCGACATCGAGCCC 1740

Qy 1468 AGTATGATCTGTGAGCAAAATCTTCAGAAACCTGGGCAAGACATGCCAGGGTAGACT 1527

Db 1741 AAGATCTGTGTGGCGGCTACT---CGACTCAAGTCCGAGCGCTGGCCAGGGGAGCTCA 1797

Qy 1528 GAGAGGCCCCCTGACCTGTGAGGAGAGAGGACACTTACTACGCTATAGGATAGTAGCTGG 1587

Db 1798 GGGGGGCCCCCTGAGCTGTGCAAGAAAGAGCGGTGCTTACCTTATCGGCAATTCAGCTGG 1857

Qy 1588 GGCCTTGAGTGTGAG-----AAGAGCGCCAGGGGCTTACACCCACAGTTACCAATTTCTGT 1641

Db 1858 GGTGACGGCTGTGGGGCGGCTCCACAGCGCGGGGGTCTTACACCGCGTGGCCAACTATGTG 1917

Qy 1642 AATTGATCAACGACCAT 1661

Db 1918 GACTGATACAGACCGGAT 1937

RESULT 7

US-09-880-107-1612

Sequence 1612, Application US/09880107

Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.

APPLICANT: Vockley, Joseph G.

APPLICANT: Schertl, Uwe

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO

CURRENT APPLICATION NUMBER: US/09/880,107

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US 60/237,054

PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1612

LENGTH: 2036

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D14012
US-09-880-107-1612

```

```

Query Match      6.88; Score 114.4; DB 10; Length 2036;
Best Local Similarity 49.6%; Pred. No. 7.2e-25;
Matches 605; Conservative 0; Mismatches 561; Indels 54; Gaps 10;

```

```

OY 469 ACCCCGACAGATGAGGCTACCTGCTCCCGGCAATAGGAGATCCAACTTCACTGT 528
DB 745 AGCCCTTGCTGCTAGACGGGAGCACTGCCACTGATGTGGCCACGGGACACCTGTGT 804
OY 529 GCGCTGCCGACCAAGTTCAGGAGAAATCTGTGAATAGTTCTGATGA---CTGCTAT 585
DB 805 GCGTGGCCACAGAGCTTGTGCGAGCGGCTCTGCACATGACCTGATGAGCGCTGCTC 864
OY 586 GTTGGGATGGCTACTTTCAGGAGAAATGATAGACATCAGACACACATGCTGC 645
DB 865 TTGGGGAGAGCGCATGCGGTACCGTGGCGTGGCCAGACCTCAGCTCGGCTGAGCTGC 924
OY 646 CTTTACTGACATCCACCTCTCTTTCAGAGAGATTTACACATGTTTATGAGATGCT 705
DB 925 CTGGCTTGAGACTCCGATCTGCTTACAGAGACTGCAGCTGACTCCGTGGCGCGCG 984
OY 706 GAAACCATGGGATTTGGGAGACACATTTCTGCAAGAACCCGATGGGAGAGAAACCC 765
DB 985 GCGCTGCTGGGCTGGGCCCCCATGCTTACTGCCGAAATCCGGAAATGACAGAGAGCC 1044
OY 766 TGTGCTTTATTAAGATTACCAATGACAAAGTGAATGGAATACTGTGATGTCTGACC 825
DB 1045 TGTGTGCTAGT---GGTGAAGACAGCGCGCTCTGCTGGAGTAATCTGCGCTGAGAGCC 1101
OY 826 TGTCTAGCCAGAGAGTGGCTTACCCAGAGAAAGCCCACTGACCATTCACCAACCTT 885
DB 1102 TGGGATTCCTT-----CACCAAGTCCACATCTCACCGGATCTCTGCGACCTTG 1152
OY 886 CCGGGTTTGTACTCTGTGGAAGACTGAGATAGCAGAGAGAGAGATCAAGAAATCTAT 945
DB 1153 CCGAGCCAGCCCTCCCGGGGCGGCGAGGCTCGGCGAGAGGACACAAAGAGAGCTTG 1212
OY 946 GGAGGCTTTAAGAGCAGGCGGAGACACCCATGAGCAGGCTCCCTCATCTCTGCTG 1005
DB 1213 CTGGGCGCACGATATATGAGGGGCTCTCTCTGCTGCTCGCGCTGACACCTG----- 1265
OY 1006 CCTGTGACCATCTCCATGCCCCAGGGCACTTGTGTGGGGCGCTGATCACCCCTGTC 1065
DB 1266 --GCTGGCGGCATCTACATCGGGGAGCAGCTTCTGCGCGGAGGCTGTCACACCTGC 1323
OY 1066 TGGGTCTCACTGTGCCCACTGACGACATTAANAACGACATCTTA-----AGGTG 1119
DB 1324 TGGGTGTGTGCGGCGGCGGCACTGTCTTCTCCACAGCCCCCGGAGGACAGGCTTCCG 1383
OY 1120 GTCTTAGGGGACCAAGACTTGAAGAAGAAATTTTCATGAGCAGAGCTTAAAGGTGCA 1179
DB 1384 GTCTTAGGGGACCACTTCTTCAACCGCAGCAGAGAGAGTACGACATCTTGGGATGAG 1443
OY 1180 AAGATTTCAATGACGCACTTACATGAATGAAGAGATGATTCGCCAATGATATGCA 1239
DB 1444 AAGTATATCCGTACACCTGTACTGTGGTGTACACCCAG---CGACCAAGACATGTGTC 1500
OY 1240 TTGCTTAAGTTAAAGCAAGTATGTGCTGTCTTAAGATTCCAATAATGAGAAACT 1299
DB 1501 CTGATCCGCTCAAGAAAGAGGAGACCGCTGTGCACACGCTGCAATGTGTGACACCC 1560
OY 1300 GTGCTGTGCTGATG-----GGTCTTCCCTCTGGGAGTGAATGCAATCTGCTGC 1353
DB 1561 ATCTGCTGCTGCGGAGCGGAGCACTTCCCGGAGAGACAAAGTGTCAATATGGCGGG 1620
OY 1354 TGGGG-----TGTTCAGAAACAGAAAGAGGTCCCGCAGCTTCGTGATGCCAAAGT 1407
DB 1621 TGGGGCACTTGTGATGAGAACGTGAGCGGCTACTCAGCTCCCTGCGGAGAGGCTGTGTC 1680

```

```

OY 1408 AAGCTATGTCGCAACACTTTGTGCACTCCCGCACTATGACCAATGATGATGAC 1467
DB 1681 CCCCTGTGCGGAGCAAGATGACAGCCCTGTAGGTTCAGGCGCGGACATCAAGCCC 1740
OY 1468 ACTATGATCTGTGCAAGAAATCTTGAAGAACTGGGCAAGACACTGCGAGGTGACTCT 1527
DB 1741 AACATGCTGTGCGCGCTACTT---CGACTGACAGTCCGACGCTGCGGAGGAGCTCA 1797
OY 1528 GGAGGCCCCCTGACTGTGGAAGAGCGGCACTTACGTCTATGAGATGAGCTG 1587
DB 1798 GGGGGGCCCTGCGCTCGAGAGAGGCGCTGAGCTTACCTTACGCAATCAAGCTGG 1857
OY 1588 GCGCTGAGTGTGAG-----AAGAGCCAGGGGTCTACACCAAGTTACCAAAATCTG 1641
DB 1858 GGTGACGGCTCGGGCGGCTCCACAAGCGGGGGGTCTACACCGCGTGGCAACTATGTG 1917
OY 1642 AATTGATCAAGCGCACCAT 1661
DB 1918 GACTGATCAAGACCGGAT 1937

```

```

RESULT 8
US-09-879-792-33
; Sequence 33, Application US/09879792
; Patent No. US2002061850A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE OF INVENTION: Prolease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 614
; TYPE: DNA
; ORGANISM: mouse
US-09-879-792-33

```

```

Query Match      4.38; Score 72.8; DB 10; Length 614;
Best Local Similarity 57.2%; Pred. No. 2.9e-12;
Matches 175; Conservative 0; Mismatches 122; Indels 9; Gaps 2;

```

```

OY 1381 TCCCGACAGCTCTGGATGCGCAAGTCAAGCTGATGCAACACTTTGTGCAACTCCGC 1440
DB 195 TCTCCCTCTCCGAGAGGTTCAGTCAACCTGATGATCTTAAGAAAGTGCATGACATC 254
OY 1441 CAACCTTATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 255 TTGGTCTATGACACTTACTTACCCCAAGAGATGATGTGTCGCGGAGATCTACAGAGA 311
OY 1501 GGGCAAGACCTGCGAGGCTGCTGAGAGCGCCCTGACCTGTGAGAAAGAGCGCAC 1560
DB 312 GGGAGGAGCTCTGCGAGGAGAGAGAGTGAAGAGACTCTGCTGTGAGCAAGAAATGTC 371
OY 1561 TACTAGCTTATGAGATGAGAGCTGGGCTTGAAGTGTG-----AGAAGAGCGCACGG 1614
DB 372 TGTGATCTGCGAGGTGTACACAGCTGGGGCACAGGCTGTGGCCAGAAAAACAAGCTGT 431
OY 1615 GTCTACACCAAGTATACCAATCTCTGATTTGATCAAGACCACTCAAAAGTGAAGT 1674
DB 432 GTCTACACCAAGTATACCAATCTCTGATTTGATCAAGAAAGTGAAGTGAAGTGA 491

```

OY 1675 GCCTTC 1680
DB 492 CGATTG 497

RESULT 9

US-09-969-271-6
; Sequence 6, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-6

Query Match 4.3%; Score 72.2; DB 10; Length 1689;
Best Local Similarity 50.5%; Pred. No. 8,4e-12;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

OY 1028 AGGGCACTTGTGTGGTGGGGGCGGTGATCCACCCCTGTGGGTGCTCAGTCGCGCCACT 1087
DB 1013 AGCGTTCTGTGTGGGGGCGCATCTCAGCTCCTGTGATTTCTCTGCGCGCCACT 1072
OY 1088 GCACCGA-----CATAAACAGACATCTAAAGTGTGTGCTAAGGGACCAAGCTGA 1141
DB 1073 GCTTCCAGAGAGAGTTCCGCCACACTGACGATGATCTTGGGCGAGAACATCCGGG 1132
OY 1142 AGAAGAAGATTTCATGACAGAGCTTAAAGGTGCGAAGATATTCAGTACAGCCACT 1201
DB 1133 TGGTCCCTGGCGAGAGAGACAAATTTGAAGTCCGAAATATCTTCCATTAAGGAT 1192
OY 1202 ACATGAAGAGATGATTTCCCATGATTTGCTCAAGTTAAAGCCAGTGG 1261
DB 1193 TCGAT-----GATGACACTTACGACATGACATTGGCTGCTGCGAGTGAATCGGATT 1246
OY 1262 ATGTGCACTGTGCTAGATCCAAATACGTGAAGACTGTGTGCTTGAATGGTCTT 1321
DB 1247 CGTCCCGCTGTGCCAGAGAGAGAGGTGTGCCGACGTGTGCTTCCCGCGGGGACC 1306
OY 1322 TTCCCT-----CTGGGAGTGAAGTGCACATCTCGCTGGGGTGTACAGAAA 1369
DB 1307 TGCAGCTCCCGAGCTGACGAGAGTGTGAGCTCTCCGCTACGGCAAGCAYGAGGCTTGT 1366
OY 1370 CAGGAAGAAGGTCCCGGACGCTCTGATGCCAAAGTCAAGTATTGCAACACTTTGT 1429
DB 1367 CTCCTTTTATTTGAGAGAGCGGCTGAAGAGGCTCATGTGACAGCTGTACCATCAGCCGT 1426
OY 1430 GCACTCCCGCAACTGTATGACACATGATTTGATGACAGTATGATCTGCGAGAAATC 1489
DB 1421 GCACATTCACAATTTACTTAACAGACAGTACCACACATCTGTGTGCTGGAGACA 1486
OY 1490 TTCAGAAACCTGG-----GCAAGACACTGTGCGAGGGTGAATCTGGAGGCG 1534
DB 1487 CTCGGAGGGGGGCGCCCGAGCAAACTTGCAGAGCGCTGACAGGCGCATTTGGGAGGCG 1546
OY 1535 CCCTGACCTGTGAAGAGAGCGACCTACTAGTCTATGGAGATAGTAGAGCTGGGCTGG 1594
DB 1547 CCCTGGTGTGTGAAGAGATGAGCGGCTGAGTGTGGTGGGATCATCAGCTGGGCGCTGG 1606
OY 1595 AGTGTGAAGAGG-----CCAGGGGTCTACACCAAGTTACCAATTCGAAATTGGA 1648
DB 1667 TTCTGTACACAT 1679

RESULT 10

US-09-969-271-5
; Sequence 5, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-5

Query Match 4.3%; Score 72.2; DB 10; Length 2519;
Best Local Similarity 50.5%; Pred. No. 1,1e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

OY 1028 AGGGCACTTGTGTGGTGGGGGCGGTGATCCACCCCTGTGGGTGCTCAGTCGCGCCACT 1087
DB 1089 AGCGTTCTGTGTGGGGGCGCATCTCAGCTCCTGTGATTTCTCTGCGCGCCACT 1148
OY 1088 GCACCGA-----CATAAACAGACATCTAAAGGTGTGCTAAGGGACCAAGCTGA 1141
DB 1149 GCTTCCAGAGAGAGTTCCGCCACACTGACGATGATCTTGGGCGAGAACATCCGGG 1208
OY 1142 AGAAGAAGATTTCATGACAGAGCTTAAAGGTGCGAAGATATTCAGTACAGCCACT 1201
DB 1209 TGGTCCCTGGCGAGAGAGACAAATTTGAAGTCCGAAATATCTTCCATTAAGGAT 1268
OY 1202 ACATGAAGAGATGATTTCCCATGATTTGCTCAAGTTAAAGCCAGTGG 1261
DB 1269 TCGAT-----GATGACACTTACGACATGACATTGGCTGCTGAGCTGAATTCGATT 1322
OY 1262 ATGTGCACTGTGCTAGATCCAAATACGTGAAGACTGTGTGCTTGAATGGTCTT 1321
DB 1323 CGTCCCGCTGTGCCAGAGAGAGAGGTGTGCCGACGTGTGCTTCCCGCGGGGACC 1382
OY 1322 TTCCCT-----CTGGGAGTGAAGTGCACATCTCTGCTGGGGTGTACAGAAA 1369
DB 1383 TGCAGCTCCCGAGCTGACGAGAGTGTGAGCTCTCCGCTACGGCAAGCAYGAGGCTTGT 1442
OY 1370 CAGGAAGAAGGTCCCGGACGCTCTGATGCCAAAGTCAAGTATTGCAACACTTTGT 1429
DB 1443 CTCCTTTTATTTGAGAGAGGCTGAAGAGGCTCATGTGACAGCTGTACCATCAGCCGT 1502
OY 1430 GCACTCCCGCAACTGTATGACACATGATTTGATGACAGTATGATCTGCGAGAAATC 1489
DB 1503 GCACATTCACAATTTACTTAACAGACAGTACCACACATCTGTGTGCTGGAGACA 1562
OY 1490 TTCAGAAACCTGG-----GCAAGACACTGTGCGAGGGTGAATCTGGAGGCG 1534
DB 1563 CTCGGAGGGGGGCGCCCGAGCAAACTTGCAGAGCGCTGACAGGCGCATTTGGGAGGCG 1622
OY 1535 CCCTGACCTGTGAAGAGAGCGACCTACTAGTCTATGGAGATAGTAGAGCTGGGCTGG 1594
DB 1623 CCCTGGTGTGTGAAGAGATGAGCGGCTGAGTGTGGTGGGATCATCAGCTGGGCGCTGG 1682
OY 1595 AGTGTGAAGAGG-----CCAGGGGTCTACACCAAGTTACCAATTCGAAATTGGA 1648

```

Db 1683 GCTGTGACAGAGATGTCCTCCGGTGTGTACACCAAGTTACCAACTACCTGGA 1742
Qy 1649 TCAAGCCACCAT 1661
Db 1743 TTCTGTGACAACT 1755

```

RESULT 11

```

US-09-974-298-144
; Sequence 144, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/228,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 144
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Inocyte ID NO. US20020156263A1 1001470CBI
; NAME/KEY: unsure
; LOCATION: 2635
; OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-144

```

```

Query Match 4.3%; Score 72.2; DB 9; Length 2641;
Best Local Similarity 50.5%; Pred. No. 1,le-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

```

```

Qy 1028 AGGGCCACTTCGTGTGGGGCGCTGATCCACCCCTGTGGGTGCTGCTACTGCTGCCACT 1087
Db 1209 AGCGGTTCTGTGGGGGCGATCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1268
Qy 1088 GCACCGA-----CATAAACACAGACATCTAAAGGTGTGCTAGGGACCGACCTGA 1141
Db 1269 GCTTCCAGAGAGGTTTCCGCCCACTGACGTGATCTTGGGAGCAACATACCGGG 1328
Qy 1142 AGAAAGAAATTTATGACGAGCTTTAGGTGCAAGAATATTCAATGACGACCT 1201
Db 1329 TGGTCCCTGGCGAGGAGGAGCAAAATTTGAAGTCGAAAAATACATGTCATTAAGAAAT 1388
Qy 1202 ACAATGAAAGATGAGATTCCCAATGATATTGCTGCTCAAGTTAAAGCAGTGG 1261
Db 1389 TCGAT-----GATGACACTTACGACATGACATTTGCGTGTGTGAGCTGAATTCGATTT 1442
Qy 1262 ATGTGATCTGTGCTTGAATTCAAATACCTGAGAAGCTGTGCTTGCCTGATGGTCT 1321
Db 1443 CGTCCGCTGTCCAGAGAGAGCGCTGTGCGCACTGTGTGCTTCCCGCGGAC 1502
Qy 1322 TTTCCT-----CTGGAGTAGTGGCCACATCTCTGTGGGTGGGTGTTACAGAAA 1369
Db 1503 TGGAGCTGCGCGATGACCGAGTGTAGTCTCCGGCTACGGGACGAGATGAGGCTTGT 1562
Qy 1370 CAGGAAAGGTCCTCCGACGCTCTGTGATGCCAAAGTCAAGCTGATTGCAACACTTTGT 1429
Db 1563 CTCTTTTCTATTGAGAGGCTGAAAGGCTCATGTGATGATGACCATTCAGCGCGCT 1622
Qy 1430 GCACTTCCGCCACTCTATGACCATGATGATGATGATGATGATGATGATGATGATGATGATG 1489
Db 1623 GCACATCACAACATTTACTTAAAGAGAGTCAAGAGATGATGATGATGATGATGATGATGATG 1682
Qy 1490 TTGAGAAACTGG-----GCAAGACACTGCGGAGGTGACTCTGGAGGCC 1534
Db 1683 CTGGAGCGCGGGCCCGGCAAACTTGTGACGAGCGCTGCCAGGGGATTTGGGAGGCC 1742

```

```

Qy 1535 CCCTGACCTGTGACAGAGCGGACCTACTAGCTCTATGGATAGTGAAGCTGGGCTTGG 1594
Db 1743 CCCTGGTGTCTGTGAACATGAGCGGATGACCTTTGGTGGGATATGACGTGGGCTTGG 1802
Qy 1595 AGTGTGAGAGAG-----CGAGGGTCTACACCAAGTTACCAAAATTCCTGAATTTGA 1648
Db 1803 GCTGTGACAGAGATGTCTCCGGGTGTGTACACCAAGTTACCAACTACTGACTGGA 1862
Qy 1649 TCAAGCCACCAT 1661
Db 1863 TTCTGTGACAACT 1875

```

RESULT 12

```

US-09-888-615-45
; Sequence 45, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: DLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-615-45

```

```

Query Match 3.9%; Score 66.4; DB 10; Length 1614;
Best Local Similarity 55.9%; Pred. No. 5,2e-10;
Matches 171; Conservative 0; Mismatches 126; Indels 9; Gaps 2;

```

```

Qy 1381 TCCCGCCAGCTCTCTGATGTCGAAGTCAAGTATGCAACACTTTGTGCAACTCCGCG 1440
Db 1300 TCCCGCTTCTCCGGAGGTGCGAGTCAATCTCATGCTCAAGAAATGCAATGACTAC 1359
Qy 1441 CAACTCTATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Db 1360 TTGTGCTATGACATTTACTTACCCCAAGATGATGTGTGTGTGTGTGTGTGTGTGTGT 1416
Qy 1501 GGGCAAGACCTTCCAGGCTGACTCTGAGGCCCCCTGACCTGTGAGAGAGCGGAC 1560
Db 1417 GCGAGAGACTCTTCCAGGAGACAGCGGGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGT 1476
Qy 1561 TACTATGCTTATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1614
Db 1477 TGTGATCTGAGAGTGTGACACAGCTGTGGGACAGAGCTGTGTGTGTGTGTGTGTGTGT 1536
Qy 1615 GTCTACACCAAGTTACCAAAATTTCTGAATTGATCAAGACCATCAAAAGTGAAGT 1674
Db 1537 GTGTACACCAAAAGTGAAGAGTGTCTTCCCTGATTTACAGCAAGATGAGAGCGAGTGT 1596
Qy 1675 GGCCTTC 1680
Db 1597 CGATTTC 1602

```

RESULT 13

```

US-09-879-792-11
; Sequence 11, Application US/09879792
; Patent No. US20020061850A1
; GENERAL INFORMATION:

```


;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084411
;; PRIOR FILING DATE: 1998-05-06

;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 3.9%, Score 66.4; DB 9; Length 2412;

Best Local Similarity 54.6%; Pred. No. 6, 7e-10; Matches 202; Conservative 0; Mismatches 156; Indels 12; Gaps 3;

QY 1310 CTGATGGGCTCTTCCCTGAGGATGAGTCCACATCTCTGAGGCTGTTACAGAAA 1369
DB 1108 CTGAAGAGACTTCCCGATGCAAAAGTGCTGAGCGTCAAGATGGGGCCACAGAG 1167
QY 1370 CAGAAAAGGCTCCCGCAGCTCTG---GATGCCAAAGTCAAGCTGATTCACAACCTT 1426
DB 1168 ATGAGGTGAGCGCTCCCTGCTGAGACACGCGCGCTTGTGATTTCAACAGAA 1227
QY 1427 TGTCAACTCCCGCAACTCTATGACACATGATTGATGACATGATCTGTCAGAA 1486
DB 1228 TGTCAACACAGGAGCGTACGAGTGCATATCTCCCTCATGCTGTGCGGGGCT 1287
QY 1487 ATCTTCAGAACTGGGAGACACCTGCGAGGTGACGTGTAGAGGCCCTGTACGTG 1546
DB 1288 ACCT---GACGGGTGCGTGGACACTGCCAGGGGAGACAGCGGGGCCCTGTGTGTC 1344
QY 1547 AGAAGAGCGCACTACTAGTCTATGGATATGAGCTGGGCGCTGAGTG-----TG 1600
DB 1345 AAGAGAGAGAGCTGTGAAGTATGAGGAGGAGCACTTGGCATTCGCGCCAGAGG 1404
QY 1601 AGAAGAGCGGAGGCTCTACACCAAGTTACCAATTCGTAATTGATGATCAAGCCACA 1660
DB 1405 TGAACAGCGCTGGGGTGTACACCTGTTCACCTTCTTGAGTGATTCACAGCAGAGA 1464
QY 1661 TCAAAAGTGA 1670
DB 1465 TGAAGAGAGA 1474

RESULT 15
US-09-978-697-68
; Sequence 68, Application US/09978697
; Patent No. US20020169284A1

GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guiney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James.
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OR INVENTION: Acids Encoding the Same
FILE REFERENCE: P26301C27
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: US/09/978,697
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545

PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30	PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05	PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-5-07	PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13	PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13	PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13	PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085697

Query Match	3.9%	Score 66.4	DB 9	Length 2412
Best Local Similarity	54.6%	Pred. No. 6.7e-10		
Matches 202	Conservative 0	Mismatches 156	Indels 12	Gaps 3
Qy	1310	CTGATGGGCTTCCTTCCTGGGAGTGGATGCGACATCTCGTGGGGGTATACAGAA	1369	
Db	1108	CTGATAGACACTCTCCGATGAAAGTGTGTGGACCTGACGATATGGGGGCCACAGAG	1167	
Qy	1370	CAGAAAAAGGCTCCGCGACGCTCTGCG--GATGCCAAAGTCAACGATTTGCGCAACTT	1426	
Db	1168	ATGGAGGTGACGGCTCCCTCTCTCTTCAACACACGGCGCCGTCCTTTGATTTTCAACAGA	1227	
Qy	1427	TGTGCACTCTCCGCCCACTCTATATGACCCACATGATTTGATGACAGTATGATCTGTGAGAA	1486	
Db	1228	TGTGCAACCAACGGGACGTGTGACGTGGGATATATCTCCCTCCATGCTCTGTGGGGGCT	1287	
Qy	1487	ATCTTCAGAAACCTGGGCAAGACACTTGCACGGGTGACTCTGGAGGCCCTCTGACCTTGTG	1546	
Db	1288	ACCT---GACGGGTGGCGTGTGACACGCTTGGACAGGGGACAGCGGGGGCCCTCGGTGTCTC	1344	

Oy	1547	AGAGAGCAGCGCACTACTACCTCTATATGGATAGTACCTGCGGCGCTTGAGATG-----TG	16000
Db	1345	AAGAGAGGAGGCTGTGCAAGTTAGTGGGAGCGCAGCGCTTTGGCATGGGCTGGCGAGAG	1404
Oy	1601	AGAGAGGCGCACGGGTGTACACCCCAAGTTACCAATTTCCPAATTTGGATCAAGGCCACCA	16600
Db	1405	TGAAACAAGCTGTGGGGTGTACACCGGTGTACCGCTTCTCGTACGATGGATCCAGGACAGA	1464
Oy	1661	TCAAAGGTCA 1670	
Db	1465	TGGAGAGAGA 1474	

```
Search completed: March 6, 2003, 23:07:10
Job time : 137.5 secs
```

10

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 10:42:18 : Search time 342.5 Seconds
(without alignments)
11066.024 Million cell updates/sec

Title: US-09-912-559-2

Perfect score: 1683

Sequence: 1 atgttgcgcagatgtctga.....aaagtgaaatgcttctaa 1683

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	1683	100.0	1683	24	AAL45697 Human blood coagul
2	1679.8	99.8	1683	24	AAL45696 Human blood coagul
3	1679.8	99.8	3008	24	ABN95170 Gene #1668 used to
4	1675	99.5	3623	21	AACT6693 Human ORFX ORF2248
5	526.4	31.3	617	21	AAA44763 Human secreted exp
6	241	14.3	397	24	AB199281 Mouse ischaemic co
7	120.2	7.1	451	22	ABA58323 Human foetal liver
8	120.2	7.1	451	22	AAK06420 Human brain expres
9	120.2	7.1	451	22	AAK32096 Human bone marrow

C 10	120.2	7.1	451	22	AA137947	Probe #663 used t
C 11	120.2	7.1	451	24	AB506863	Human genome-deriv
C 12	117	7.0	117	22	ABA70919	Human foetal liver
C 13	117	7.0	117	22	AAK19188	Human brain expres
C 14	117	7.0	117	22	AAK45153	Human bone marrow
C 15	117	7.0	117	22	AAK45106	Human genome-deriv
C 16	117	7.0	117	24	AB519403	Probe #1972 used
17	114.4	6.8	2033	15	AA063951	Human genome-deriv
18	114.4	6.8	2036	24	ABN95114	Hepatocyte growth
19	114.4	6.8	2036	24	ABN95114	Gene #1612 used to
20	93	5.5	970	15	AA063945	lung cancer relate
21	87.6	5.2	2252	11	AA005673	Gene comprising HG
22	87.4	5.2	2512	12	AA012867	v-PA.alpha1. Desm
23	86.8	5.2	1323	8	AAAT0992	DM-1229 cell line
24	85.2	5.1	1422	8	AAAT0991	Modified tissue pl
25	84.8	5.0	1137	11	AA005675	v-PA.alpha2. Desm
26	79.6	4.7	2257	11	AA005674	v-PA.alpha2. Desm
27	78	4.6	1620	11	AA000543	Vampire bat plasmi
28	77.2	4.6	919	18	AAAT90048	Chicken urukinase
29	75.4	4.5	1689	10	AAAT91217	Mutated CDNA codin
30	73.8	4.4	2097	12	AAQ12871	T-PA with -ve char
31	73.8	4.4	2100	12	AAQ12874	T-PA variant havin
32	72.8	4.3	614	24	AAQ27741	Human full length
33	72.2	4.3	329	15	AAQ79004	Gene comprising HG
34	72.2	4.3	1047	14	AAQ40658	tpa-2 CDNA. Synth
35	72.2	4.3	1065	18	AAAT87456	mt-PA6-E. Homo sa
36	72.2	4.3	1065	18	AAAT62606	DNA encoding t-PA
37	72.2	4.3	1068	14	AAO53322	Human truncated tP
38	72.2	4.3	1068	14	AAO53322	Human truncated tP
39	72.2	4.3	1068	14	AAO40657	tpa-1 CDNA. Synth
40	72.2	4.3	1068	14	AAO40660	tpa-6 CDNA. Synth
41	72.2	4.3	1068	14	AAO40667	tpa-11 CDNA. Synth
42	72.2	4.3	1068	14	AAO40669	tpa-12 CDNA. Synth
43	72.2	4.3	1068	14	AAO40671	tpa-13 CDNA. Synth
44	72.2	4.3	1068	14	AAO40673	tpa-14 CDNA. Synth
45	72.2	4.3	1068	14	AAO40677	tpa-16 CDNA. Synth

ALIGNMENTS

RESULT 1	
AAL45697	standard; DNA; 1683 BP.
ID	
AC	AAL45697;
XX	
DT	13-JUN-2002 (first entry)
XX	
DE	Human blood coagulation factor VII activating protease mutant DNA.
XX	
KW	Human: blood coagulation factor VII activating protease; FSAP.
KW	Single-chain plasminogen activator; bleeding disorder; haematological;
KW	haemostatic; mutant; gene; ds.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
FH	Key
FT	CDS
FT	1..1683
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "mutant FSAP"
PN	EP1182258-A1.
XX	
PD	27-FEB-2002.
XX	
PF	05-JUL-2001; 2001EP-0115691.
XX	
PR	26-JUL-2000; 2000DE-1036641.
PR	10-OCT-2000; 2000DE-1050040.
PR	21-OCT-2000; 2000DE-1052319.
PR	12-APR-2001; 2001DE-1018706.

RESULT 2
AAL45696
ID AAL45696 standard: DNA; 1683 BP.
XX
AC AAL45696;
XX
D7 13-JUN-2002 (first entry)
XX
DE Human blood coagulation factor VII activating protease DNA.
XX
KW Human; blood coagulation factor VII activating protease; FSAP;
KW Single-chain plasminogen activator; bleeding disorder; haematological;
KW haemostatic; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1683
FT /tag= a
FT /product= "FSAP"
XX
XX EP1182258-A1.
XX
PD 27-FEB-2002.
XX
PF 05-JUL-2001; 2001EP-0115691.
XX
PR 26-JUL-2000; 2000DE-1036641.
PR 10-OCT-2000; 2000DE-1050040.
PR 21-OCT-2000; 2000DE-1052319.
PR 12-APR-2001; 2001DE-1018706.
XX
XX (AVENTIS BEHRING GMBH.
PI Roemisch J, Steehr H, Feussner A, Lang W, Weimer T, Becker M;
PI Nerlich C, Muth-Naumann G;
XX
DR MPI; 2002-270939/32.
DR P-PSDB: AAO17144.
XX
PT New nucleic acid encoding mutant factor 7 activating protease, useful
PT for diagnosis, treatment and prevention of coagulation disorders, also
PT related protein and antibodies
XX
PS Claim 2: Page 15-16; 27pp; German.
XX
XX The present invention relates to a mutant of the DNA sequence encoding
CC the protease (FSAP) that activates blood coagulation factor VII (FVII)
CC and single-chain plasminogen activator, where at least one of the base
CC changes G to C at nucleotide 1177 and G to A at nucleotide 1601 is
CC present. The mutant sequences can be used in the treatment and prevention
CC of bleeding disorders associated with inherited or acquired defects of
CC blood coagulation factors V, VIII, IX, X, XI, XII, von Willebrand's
CC factor and/or with antibodies against any of these proteins. The present
CC sequence is the human FSAP coding sequence.
XX
SO Sequence 1683 BP; 440 A; 436 C; 437 G; 370 T; 0 other;
Query Match 99.8%; Score 1679.8; DB 24; Length 1683;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 121 CAGTATGATTACAGGCTACGAGATTATATACAGGAAGAGAACACCAGTACACTTACC 180
Qy 181 CATGCTGAGAAATCTGACTGGTACTACACTGAGACCAGTGCATCTCATGCCCAAC 240
Db 181 CATGCTGAGAAATCTGACTGGTACTACACTGAGACCAGTGCATCTCATGCCCAAC 240
Qy 241 CCCTGTGAACACGCTGGGACACTGCTCCATGAGAGCACCTTCACATGACAGTGCCTG 300
Db 241 CCCTGTGAACACGCTGGGACACTGCTCCATGAGAGCACCTTCACATGACAGTGCCTG 300
Qy 301 GCTCCTTTCCTGCGAATTAAGTTCAGAAAGTGCAAAATATGATGCAAGACCAACCTAT 360
Db 301 GCTCCTTTCCTGCGAATTAAGTTCAGAAAGTGCAAAATATGATGCAAGACCAACCTAT 360
Qy 361 GGCCGGGGCCCAATGCTCATTAACCCAGATCTCCCTACTACCGCTGTGTCTGTAACAC 420
Db 361 GGCCGGGGCCCAATGCTCATTAACCCAGATCTCCCTACTACCGCTGTGTCTGTAACAC 420
Qy 421 CATTACACAGTCCCGACGCTCCCAAGTGGTTCCTGTATGAGAGGCCCAACCCCTGCGAC 480
Db 421 CATTACACAGTCCCGACGCTCCCAAGTGGTTCCTGTATGAGAGGCCCAACCCCTGCGAC 480
Qy 481 AATGGGGCTACCTGCTCCCGCATTAAGCGAGATCCAGTTGACCTGTGCTGCCGAC 540
Db 481 AATGGGGCTACCTGCTCCCGCATTAAGCGAGATCCAGTTGACCTGTGCTGCCGAC 540
Qy 541 CAGTTCAAGGGGAATTTCTGTGAATTAAGTTCTGATGACCTGATGTTGGGATGGCTAC 600
Db 541 CAGTTCAAGGGGAATTTCTGTGAATTAAGTTCTGATGACCTGATGTTGGGATGGCTAC 600
Qy 601 TCTTACCGAGGGAATTAAGTATGACATCAACAGCATGCGCTTACTGGAACCTCC 660
Db 601 TCTTACCGAGGGAATTAAGTATGACATCAACAGCATGCGCTTACTGGAACCTCC 660
Qy 661 CACCTCCTCTGAGAGAGATTTACACATGTTATGAGATGCTGGAACCCATGGGATT 720
Db 661 CACCTCCTCTGAGAGAGATTTACACATGTTATGAGATGCTGGAACCCATGGGATT 720
Qy 721 GGGGACACCAATTTCTGAGAAACCCAGATCGGACGAAAAAGCCCTGCTGTTATTTAA 780
Db 721 GGGGACACCAATTTCTGAGAAACCCAGATCGGACGAAAAAGCCCTGCTGTTATTTAA 780
Qy 781 GTTACCAATGACAGAGTGAATGGAATACGTGATGCTCAGCGCTGCTCAGCCAGAC 840
Db 781 GTTACCAATGACAGAGTGAATGGAATACGTGATGCTCAGCGCTGCTCAGCCAGAC 840
Qy 841 GTTGTCCATCCACAGAGGAAGCCCATGAGCATCAACCAAGCTTCCGGGTTTGACTCC 900
Db 841 GTTGTCCATCCACAGAGGAAGCCCATGAGCATCAACCAAGCTTCCGGGTTTGACTCC 900
Qy 901 TGTGGAAGACTAGATATACAGAGGAAGATCAAGAGATCTATGAGAGCTTTAAAGAC 960
Db 901 TGTGGAAGACTAGATATACAGAGGAAGATCAAGAGATCTATGAGAGCTTTAAAGAC 960
Qy 961 ACGGCGGGCAAGCACCATGAGGAGCGCTCCCTCAGTCTCGCTGCTCTGACATCTCC 1020
Db 961 ACGGCGGGCAAGCACCATGAGGAGCGCTCCCTCAGTCTCGCTGCTCTGACATCTCC 1020
Qy 1021 ATGCCCCAGGGCCATCTCTGTGTGGGCGCTGATCCACCCCTGCTGGGTCTCATGCT 1080
Db 1021 ATGCCCCAGGGCCATCTCTGTGTGGGCGCTGATCCACCCCTGCTGGGTCTCATGCT 1080
Qy 1081 GCCCATGCAACCCAGATTAATAAACCCAGACATTAAGGTGTGCTAGAGGACAGACCTG 1140
Db 1081 GCCCATGCAACCCAGATTAATAAACCCAGACATTAAGGTGTGCTAGAGGACAGACCTG 1140
Qy 1141 AAGGAAGAAGAAATTTATAGAGCAGAGCTTTAGGGTCAAGAAGATATTCAGATACGCCAC 1200
Db 1141 AAGGAAGAAGAAATTTATAGAGCAGAGCTTTAGGGTCAAGAAGATATTCAGATACGCCAC 1200
Qy 1201 TACATATCAAGACAGATGAGATTCCTCCCAATGATATTCATTCCTCAAGTTAAAGCAGTG 1260
Db 1201 TACATATCAAGACAGATGAGATTCCTCCCAATGATATTCATTCCTCAAGTTAAAGCAGTG 1260

Dd	1201	TACAATGAAGAGATGAGATATCTCCCAAAATGATATTGCATTGGCTCAAAGTTAAAGCAGTCG	1260
Oy	1261	GATGGTCACTGTGGCTTAGAATCCAATAACGTGAAGACTGTGTGCTTGCGCTGATGGGTCC	1320
Dd	1261	GATGGTCACTGTGCTCTTAGATCCAAATACGTGAAGACTGTGTGCTTGCGCTGATGGGTCC	1320
Oy	1321	TTTTCCCTGGAGTAGTAGTGCCACCATCTCTGGCTGGGTGTTCACGAACACAGAAAAAGG	1380
Dd	1321	TTTTCCCTGGAGTAGTAGTGCCACCATCTCTGGCTGGGTGTTCACGAACACAGAAAAAGG	1380
Oy	1381	TCCCGCACCTCTCGGATGAGCAAGTCAAGCTGATGGCAACCTTTGTGCCAACCTCCCG	1440
Dd	1381	TCCCGCACCTCTCGGATGAGCAAGTCAAGCTGATGGCAACCTTTGTGCCAACCTCCCG	1440
Oy	1441	CAACTCTATGACCACATGATTGATGACAGTATGATCTGTGCAGAAAATCTTCAGAACCCT	1500
Dd	1441	CAACTCTATGACCACATGATTGATGACAGTATGATCTGTGCAGAAAATCTTCAGAACCCT	1500
Oy	1501	GGGCAAGACACCTGCGCAGGGGTGACTGTGAGCCCCCGGACCTGTGGAAGAGAGGCGACC	1560
Dd	1501	GGGCAAGACACCTGCGCAGGGGTGACTGTGAGCCCCCGGACCTGTGGAAGAGAGGCGACC	1560
Oy	1561	TACTAGCTCTATGAGATGATGAGCTGGGCCCTGAGATGTGAGAAGAGGCGCAGGGTCTAC	1620
Dd	1561	TACTAGCTCTATGAGATGATGAGCTGGGCCCTGAGATGTGAGAAGAGGCGCAGGGTCTAC	1620
Oy	1621	ACCAAGTATACCAATTCCTGAAATTGGATCAAGCGCACATCAAAAAGTGAAGTGGCTTC	1680
Dd	1621	ACCAAGTATACCAATTCCTGAAATTGGATCAAGCGCACATCAAAAAGTGAAGTGGCTTC	1680
Oy	1681	TAA 1683	
Dd	1681	TAA 1683	
RESULT 3			
ID	ABN95170	standard; DNA; 3008 bp.	
XX	AC	ABN95170;	
XX	DE	13-AUG-2002 (first entry)	
XX	DE	Gene #1668 used to diagnose liver cancer.	
KM	KM	Genes: liver cancer; ds; hepatocellular carcinoma; hepatotropic;	
KW	KW	metastatic liver tumour; cytostatic; expression profile; disease state;	
OS	OS	disease progression; drug toxicity; drug efficacy; drug metabolism.	
OS	OS	Homo sapiens.	
PN	PN	WO200229103-A2.	
PD	PD	11-APR-2002.	
PF	PF	02-OCT-2001; 2001MO-US30589.	
PR	PR	02-OCT-2000; 2000US-237054P.	
PA	PA	(GENE-) GENE LOGIC INC.	
PI	PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;	
DR	DR	WPI; 2002-426119/45.	
PT	PT	Diagnosing and detecting the progression of liver cancer,	
PT	PT	hepatocellular carcinoma or metastatic liver tumor in a patient,	
PT	PT	involves detecting the level of expression of two or more genes in a	
XX	XX	liver tissue sample	
PS	PS	Claim 1; SEQ ID NO 1668; 298pp; English.	
XC	XC	The invention relates to a novel method for diagnosing and detecting th	

progression of liver cancer, hepatocellular carcinoma or metastatic liver
tumor in a patient, and differentiating metastatic liver cancer from
hepatocellular carcinoma in a patient, involving detecting the level of
expression of two or more genes represented in ABN93503-ABN9455 in a
tissue sample. The method of the invention has hepatotropic, and
cytostatic activity. The method is useful for diagnosing and detecting
the progression of liver cancer, hepatocellular carcinoma and metastatic
liver carcinoma in a patient. The method is useful for identifying
expression profiles which serve as useful diagnostic markers as well as
markers that can be used to monitor disease status, disease progression,
drug toxicity, drug efficacy and drug metabolism.

Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at http://wipo.int/pub/published_pat_sequences.

Query Match	99.8%;	Score 1679.8;	DB 24;	length 3008;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1681; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

QY	1	ATGTTTCCAGAGATGTCATCATCTCAATGTCGTGTTAAATGGCTGCGTGGGAAGACA	60
Db	97	ATGTTTCCAGAGATGTCATCATCTCAATGTCGTGTTAAATGGCTGCGTGGGAAGACA	156
QY	61	GCCTGTGGGTTCTCCCTATGTCTTTATTGGAAAGCTGGACCCAGACTGGACCCCTGAC	120
Db	157	GCCTGTGGGTTCTCCCTATGTCTTTATTGGAAAGCTGGACCCAGACTGGACCCCTGAC	216
QY	121	CAGTATGATATACAGCTACGACGATTATTAATCAGGAAGAACACCAAGTAGACACTTACC	180
Db	217	CAGTATGATATACAGCTACGACGATTATTAATCAGGAAGAACACCAAGTAGACACTTACC	276
QY	181	CATGCTAGATTCCTGACTGTACTACTAGAGACCAAGCTGATCCATGCCAGCCCAAC	240
Db	277	CATGCTAGATTCCTGACTGTACTACTAGAGACCAAGCTGATCCATGCCAGCCCAAC	336
QY	241	CCCTGTGAACACGCTGGGGACATGCTGTGTCATATGGAGACACTTCCATGACAGCTGCCG	300
Db	337	CCCTGTGAACACGCTGGGGACATGCTGTGTCATATGGAGACACTTCCATGACAGCTGCCG	366
QY	301	GCTCCTTTCCTGTGGAAATTAAGTGTCAAGAAAGTGCAAAATACGTGACAGGACACCATATG	360
Db	397	GCTCCTTTCCTGTGGAAATTAAGTGTCAAGAAAGTGCAAAATACGTGACAGGACACCATATG	456
QY	361	GAGCGGGGGCAATGTCCTATTACCCAGAGACCTGCCCTACTAACCGGTGTCGTATTAAC	420
Db	457	GAGCGGGGGCAATGTCCTATTACCCAGAGACCTGCCCTACTAACCGGTGTCGTATTAAC	516
QY	421	CCTTACACAGGTCACCAGCTGCTCCCAAGTGCTTCTGTATGACAGGGCAAAACCCCTGCCAG	480
Db	517	CCTTACACAGGTCACCAGGCTGCTCCCAAGTGCTTCTGTATGACAGGGCAAAACCCCTGCCAG	576
QY	481	AATGGGCACTACCGTCGCCCGGCAATAGGCGGAATCCAAAGTTCACCTGTGCTGTGCCGAC	540
Db	577	AATGGGCACTACCGTCGCCCGGCAATAGGCGGAATCCAAAGTTCACCTGTGCTGTGCCGAC	636
QY	541	CAGTTCAAGGGGAAATTCGTGAATAATAGGTCCTGTGACATGCTATGTTGGCGATGGCTAC	600
Db	637	CAGTTCAAGGGGAAATTCGTGAATAATAGGTCCTGTGACATGCTATGTTGGCGATGGCTAC	696
QY	601	TCTTACCGAGGGGAAATTAATAGGACAGTCAACAGCATGCGTGGCTTTACTGGAACTCC	660
Db	697	TCTTACCGAGGGGAAATTAATAGGACAGTCAACAGCATGCGTGGCTTTACTGGAACTCC	756
QY	661	CACCTCCTCTTGACAGGAATATCAACATGTTTATGAGAGATCTCAAAACCCATGGGATT	720
Db	757	CACCTCCTCTTGACAGGAATATCAACATGTTTATGAGAGATCTCAAAACCCATGGGATT	816
QY	721	GGGGAACACAATTTCTGACAGAAACCCAGATGGGACGCAAAAAGCCCTGGTCTTTATTAA	780
Db	817	GGGGAACACAATTTCTGACAGAAACCCAGATGGGACGCAAAAAGCCCTGGTCTTTATTAA	876

QY	781	GTTCACAAATGACAAAGSTGAATATGGGAATGACTGATGATCTCAACCTGTGTCAAGCCACAGAC	840
Db	877	GTTACCAATGACAAAGSTGAATATGGGAATGACTGATGATCTCAAGCTGTGTCAAGCCACAGAC	936
QY	841	GTTTGCTACCCAGAGAAAGCCCTGAGCCATCAACCAAGCTCTCCGGGTGTTGACTCC	900
Db	937	GTTTGCTACCCAGAGAAAGCCCTGAGCCATCAACCAAGCTCTCCGGGTGTTGACTCC	996
QY	901	TGTGGAAAGACTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGGAGCTTTAAGAGC	960
Db	997	TGTGGAAAGACTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGGAGGCTTTAAGAGC	1056
QY	961	ACGGCGGCGCAAGCACCACATGCGAGGCGTCCCTCAAGTCTCCGCTCTGACCATCTCC	1020
Db	1057	ACGGCGGCGCAAGCACCACATGCGAGGCGTCCCTCAAGTCTCCGCTCTGACCATCTCC	1116
QY	1021	ATGCCCCAGGGCAGCTTGTGTGGTGGGGCGCTGATCAACCCCTGTGGGTGTCACCTGCT	1088
Db	1117	ATGCCCCAGGGCAGCTTGTGTGGTGGGGCGCTGATCAACCCCTGTGGGTGTCACCTGCT	1178
QY	1081	GCCCCACTGCACCGACATAAAAACACACATCTAAAGGTGTGCTAGGGGACAGAGACTG	1144
Db	1177	GCCCCACTGCACCGACATAAAAACACACATCTAAAGGTGTGCTAGGGGACAGAGACTG	1236
QY	1141	AAGAAAGAGATTTTCATGACACAGCTTTAGGGTGGCAAGAGATATTCACAGTACACCCAC	1200
Db	1237	AAGAAAGAGATTTTCATGACACAGCTTTAGGGTGGCAAGAGATATTCACAGTACACCCAC	1296
QY	1201	TACAAATGAAAGATGAGATTTCCCCACAAATGATATTTGCAATTTGCTCAAGTTAAAGCCAGTG	1260
Db	1297	TACAAATGAAAGATGAGATTTCCCCACAAATGATATTTGCAATTTGCTCAAGTTAAAGCCAGTG	1356
QY	1261	GATGGTCACTGTGCTGTGAAATACGAAATACGGAAGACGTGTGCTTGGCTCGATGGGTCC	1328
Db	1357	GATGGTCACTGTGCTGTGAAATACGAAATACGGAAGACGTGTGCTTGGCTCGATGGGTCC	1418
QY	1321	TTTTCTCTGGGATGAGTGCCACATCTCTGGCTGGGCTGTTACACAAACAGGAAAGAGG	1380
Db	1417	TTTTCTCTGGGATGAGTGCCACATCTCTGGCTGGGCTGTTACACAAACAGGAAAGAGG	1476
QY	1381	TCCCGCAGACTCTGTGATGCCAAAGCTCAAGCTGATTTGGCAACACTTTGTGCAACTCCGC	1444
Db	1477	TCCCGCAGACTCTGTGATGCCAAAGCTCAAGCTGATTTGGCAACACTTTGTGCAACTCCGC	1536
QY	1441	CAACTCTATGACACACATGATTTGATGACAGTATGATCTGTGAGGAAATCTTCAGAAACCT	1500
Db	1537	CAACTCTATGACACACATGATTTGATGACAGTATGATCTGTGAGGAAATCTTCAGAAACCT	1596
QY	1501	GGGCAAGACACTGCGCAGGGGAGACTCTGGAGGCCCTCCGACTGTGAGAGAGCGGCACAC	1560
Db	1597	GGGCAAGACACTGCGCAGGGGAGACTCTGGAGGCCCTCCGACTGTGAGAGAGCGGCACAC	1656
QY	1561	TACTACGTCGTATGGGATAGTGAGCTGGGGCTGTGAGTGTGAAGAGAGCCAGGGGTCTAC	1620
Db	1657	TACTACGTCGTATGGGATAGTGAGCTGGGGCTGTGAGTGTGAAGAGAGCCAGGGGTCTAC	1718
QY	1621	ACCAAGTATACAAATTCCTGGAATTTGGATTCAAAGCCACATCAAAATGTAATGGCTTC	1680
Db	1717	ACCAAGTATACAAATTCCTGGAATTTGGATTCAAAGCCACATCAAAATGTAATGGCTTC	1776
QY	1681	TAA 1683	
Db	1777	TAA 1779	
RESULT 4			
AACT6693			
ID	AACT6693	standard, cDNA, 3623 BP.	
AC	AACT6693:		
XX			
XT	08-FEB-2001	(first entry)	

DE Human ORFX ORE248 polynucleotide sequence SEQ ID NO:4495.

KW Human: open reading frame; ORFX: detection; cytosolic; hepatotropic;
KW vlnnerary; antipsoiatric; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antirheumatic; immunosuppressant; cardiant;
KW hypotenostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW immunostimulant; dermatological; immunosuppressive; antineoplastic;
KW antiviral; antibacterial; antifungal; antineumatic; antihypoid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.

OS Homo sapiens.

XX
XX
XX WC200058473-AZ.

PD 05-OCT-2000.

XX
XX
XX 31-MAR-2000; 2000WO-US08621.

PF 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.

XX
XX (CURA-) CURAGEN CORP.

PA
XX
XX Shimkels RA, Leach M;

PI
XX
XX WPI: 2000-602362/57.
DR P-PSDB: AAB42484.

XX
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX

PS Claim 5; Page 3679-3681; 5507pp: English.

XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vlnnerary;
CC antipsoiatric; antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antirheumatic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotenostimulant; dermatological; immunosuppressive;
CC antineoplastic; antineumatic; antihypoid; antifungal; antineumatic;
CC antihypoid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antineoplastic disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX
XX Sequence 3623 BP; 848 A; 1005 C; 930 G; 840 T; 0 other;

Query Match 99.5%; Score 1675; DB 21; Length 3623;
Best Local Similarity 99.7%; Pred. No. 0;
Mismatch 1678; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 ATGTTTGACAGATGCTGATCCAGATGTTGCTGTATATGCTGTGCGAAGACCA GCG

Db	97	ATGTTTGCCAGAAATGTCTGAAATCTCCATGTTCTCTGTTAAATGGGTCTGGTGGGAAAGACA	156
Oy	61	GCGTGGGGTTCCTCCGTGATGTCTTTATTGAAAGCCTGGACCCAGACATGACCCCTTAC	120
Db	157	GCGTGGGGTTCCTCCGTGATGTCTTTATTGAAAGCCTGGACCCAGACATGACCCCTTAC	216
Oy	121	CAGTATGATTACAGTACGAGAGATTATATCAGAGAAAGAACACAGTATGACACTTACC	180
Db	217	CAGTATGATTACAGTACGAGAGATTATATCAGAGAAAGAACACAGTATGACACTTACC	276
Oy	181	CATGCTGAGAAATCCTGACTGGTACTACACTGAGAGACCAGCTGATCCATGACGCCAAC	240
Db	277	CAGCGTGAATAATCCGTAGCTGTACTACACTGAGAGACCAGCTGATCCATGAGGCCAAC	336
Oy	241	CCCTGTGAACAGGGTGGGAGCTGCCTCGTCCATGSGGACACTTCAATGACAGCTCGTG	300
Db	337	CCCTGTGAACAGGGTGGGAGCTGCCTCGTCCATGSGGACACTTCAATGACAGCTCGTG	396
Oy	301	GCTTCCTTTCTCTGGGSAATTAAGTGTCAAGAAAGTGCAGAAATACGTGCAAGACACCACTGT	360
Db	397	GCTTCCTTTCTCTGGGSAATTAAGTGTCAAGAAAGTGCAGAAATACGTGCAAGACACCACTGT	456
Oy	361	GAGCGGGGCAATGTCTCATCTATACCAGAGACTCCTCCCTACTACCGCTGTGCTGTAAACAC	420
Db	457	GAGCGGGGCAATGTCTCATCTATACCAGAGACTCCTCCCTACTACCGCTGTGCTGTAAACAC	516
Oy	421	CCTTACACAGGTCACAGTGTCTCCCAAGTGTCTCTGTATGACAGCCAAACCCCTGGCAG	480
Db	517	CCTTACACAGGTCACAGTGTCTCCCAAGTGTCTCTGTATGACAGCCAAACCCCTGGCAG	576
Oy	481	AATGGGGTCACTGCTCCCGGCAATAGCGGSAANTCCAAGTTCACTGTGCTGTCCCGAC	540
Db	577	AATGGGGTCACTGCTCCCGGCAATAGCGGSAANTCCAAGTTCACTGTGCTGTCCCGAC	636
Oy	541	CAGTCAAGGGSAATTTCTGTGAAATTAAGTGTGATGTGACTCTGTGTGGCATAAGCTAC	600
Db	637	CAGTCAAGGGSAATTTCTGTGAAATTAAGTGTGATGTGACTCTGTGTGGCATAAGCTAC	696
Oy	601	TCTTACCAGGGSAATAATAGTAGAGACAGTCAACCAAGCTGCTCTTATCTGGAATCC	660
Db	697	TCTTACCAGGGSAATAATAGTAGAGACAGTCAACCAAGCTGCTCTTATCTGGAATCC	756
Oy	661	CACCTCCTCTTCAGAGAAATTAACAATGTTTATAGAGAGATGCTGAAACCCTGGGAAAT	720
Db	757	CACCTCCTCTTCAGAGAAATTAACAATGTTTATAGAGAGATGCTGAAACCCTGGGAAAT	816
Oy	721	GGGGAACACAATTTCTGCAAGAACCCAGATGCGGACGAGAAAGCCTTGCTTATTTAA	780
Db	817	GGGGAACACAATTTCTGCAAGAACCCAGATGCGGACGAGAAAGCCTTGCTTATTTAA	876
Oy	781	GTTTACCAATGSAAGTCAATGGAATGGAATGCTGATGTCGTAGCGTCTCAGCCAGGAC	840
Db	877	GTTTACCAATGSAAGTCAATGGAATGGAATGCTGATGTCGTAGCGTCTCAGCCAGGAC	936
Oy	841	GTTTCTCAACCAGAGAAAGCCCACTGAGCATCAACCAAGCTTCGGGGTTTGACTCC	900
Db	937	GTTTCTCAACCAGAGAAAGCCCACTGAGCATCAACCAAGCTTCGGGGTTTGACTCC	996
Oy	901	TGTGGAAGACTGAGATAGCAGAGAGAGAGATCAAGAGAAATCTATGAGGCTTTAAAGAC	960
Db	997	TGTGGAAGACTGAGATAGCAGAGAGAGAGATCAAGAGAAATCTATGAGGCTTTAAAGAC	1056
Oy	961	ACGGCGGGCAAGCAATGCGACGCGCTCCCACTGCTGCTGCTGTGACATCTCC	1020
Db	1057	ACGGCGGGCAAGCAATGCGACGCGCTCCCACTGCTGCTGCTGTGACATCTCC	1116
Oy	1021	ATGCGCCAGGGCCACTTCTGTGTGGGGCCGTGATGCAACCCCTCTGGGTGCTCACTGCT	1080
Db	1117	ATGCGCCAGGGCCACTTCTGTGTGGGGCCGTGATGCAACCCCTCTGGGTGCTCACTGCT	1176
Oy	1081	GCCCACTGCACGACATAAAAAACCAAGACATTTAAAGGTGTGTAGGGACACAGGACTCG	1140
Db	1177	GCCCACTGCACGACATAAAAAACCAAGACATTTAAAGGTGTGTAGGGACACAGGACTCG	1236

Oy	1141	AACAAAGAAATTTCTGAGACACTTACGGTCGAAGAATATTCAAGTACGCC	1290
Oy	1237	AGAGAAAGAGATTTTCATGACGACACTTTAGGGTGAGAGAAATTTCAAGTACGCC	1296
Oy	1201	TACATGAAGAGATGAGATTCCTCCACAATGATATTTGCATGTCTCAAGTAAAGCAGTG	1260
Db	1297	TACATGAAGAGATGAGATTTCCCCACAATGATATTTGCATGTCTCAAGTAAAGCAGTG	1356
Oy	1261	GATGCGCATGTGGCTAGAGAACCAATACGTGAAGACCTGTGCTTGCCTGATGGTCC	1320
Db	1357	GATGCGCATGTGCTTAGAATCCAAATACGTGAAGACCTGTGCTTGCCTGATGGTCC	1416
Oy	1321	TTTTCCCTCTGGGAGTGAAGTCCACATCTCTGGCTGGGGTGTTCACAGAAACAGSAAAAGG	1380
Db	1417	TTTTCCCTCTGGGAGTGAAGTCCACATCTCTGGCTGGGGTGTTCACAGAAACAGSAAAAGG	1476
Oy	1381	TCCGCGCACCTCTTGAGATGACCAAGTCAAGCTGATTCGCAACACTTTGTGTGCAACTCCGC	1440
Db	1477	TCCGCGCACCTCTTGAGATGACCAAGTCAAGCTGATTCGCAACACTTTGTGTGCAACTCCGC	1536
Oy	1441	CACCTGTATGACACATGATGATGAGACATGATCTGTGCAGSAAATCTTCAGAACTT	1500
Db	1537	CACCTGTATGACACATGATGATGAGACATGATCTGTGCAGSAAATCTTCAGAACTT	1596
Oy	1501	GGGCAAGACACCCTGCGACGGGTGACTCTGTGAGGCCCCCTGCATCTGTGAGAAAGCAGGACC	1560
Db	1597	GGGCAAGACACCCTGCGACGGGTGACTCTGTGAGGCCCCCTGCATCTGTGAGAAAGCAGGACC	1656
Oy	1561	TATACGTGTATGGGATGATGAGAGTGGGGCTGGGAGTGTGAGAAAGCAGGAGGTATC	1620
Db	1657	TATACGTGTATGGGATGATGAGAGTGGGGCTGGGAGTGTGAGAAAGCAGGAGGTATC	1716
Oy	1621	ACCCAAGTACCAAATTCCTGCAATTTGGATCAAAAGCCACCATCAAAAGTGAAGTGGCTTC	1680
Db	1717	ACCCAAGTACCAAATTCCTGCAATTTGGATCAAAAGCCACCATCAAAAGTGAAGTGGCTTC	1776
Oy	1681	TAA 1683	
Db	1777	TAA 1779	
RESULT 5			
AAAA4763			
ID	AAAA4763 standard; cDNA; 617 BP.		
XX			
AC	AAAA4763;		
XX			
DT	21-AUG-2000 (first entry)		
DE			
XX	Human secreted expressed sequence tag SEQ ID NO:1338.		
XX			
KM	Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;		
KW	expressed sequence tag; EST; probe; chemotactic; proliferative;		
KW	immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;		
KW	thrombolytic; antinflammatory; cytoskeletal; antibacterial; antifungal;		
KW	antiviral; antidiabetic; antisthmatic; vulnery; antiparkinsonian;		
KW	antitumor; osteoprotective; neuroprotective; nootropic; antipsoriatic;		
KW	cardioprotective; anticoagulant; antidepressant; gene therapy;		
KV	vaccine; autoimmune disorder; multiple sclerosis; allergic condition;		
KM	insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;		
KM	lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;		
KW	central nervous system disorder; Alzheimer's disease; stroke;		
KW	Parkinson's disease; Huntington's disease; coagulation disorder;		
KW	hemophilia; thrombosis; inflammatory disorder; Crohn's disease;		
KW	tumour; infection; depression; psoriasis; ss.		
XX			
OS	Homo sapiens.		
XX			
FN	WO200021991-A1.		
XX			
PD	20-Apr-2000.		

DE	Human secreted expressed sequence tag SEQ ID NO:1338.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

```
RESULT 5
AAAA4763
ID AAA44763 standard; cDNA; 617 BP
```

DT 21-AUG-2000 (first entry)

KW Human; mouse; chicken; rat; secreted expressed sequence tag; ssRF;
KW expressed sequence tag; EST; probe; chemotactic; proliferative;
KW Immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antitumoral;
KW antiviral; antidiabetic; antistaphylococcal; uroteric; antiparkinsonian;
KW antihcic; osteopathic; neuroprotective; nootropic; antipsychotic;
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW central nervous system disorder; Alzheimer's disease; stroke;
KW Parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW tumour; infection; depression; psoriasis; ss.

OS	Homo sapiens.
XX	
PN	W0200021991-A1
XX	
PD	20-APR-2000.

PF 15-OCT-1999; 99MO-US24206.
XX
PR 15-OCT-1998; 98US-0104436.
XX
PA (GEMV) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Metberg D, Treacy M, Bowman MR;
XX
DR WPI: 2000-317938/27.
XX
PT Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (seSTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders -
XX
PS Claim 1; Page 549; 803pp; English.
XX
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
CC sequence tags (seSTs), isolated from human, mouse, chicken and rat
CC tissue sources. The seSTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemotactic; proliferative; immunomodulatory; hematopoietic;
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytoskeletal; antibacterial; antifungal; antiviral; antidiabetic;
CC antiasthmatic; vulnereary; antineuritic; osteoprotective;
CC neuroprotective; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antidepressant. The seSTs can be used for gene
CC therapy and in vaccines. The seSTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the seSTs. Proteins encoded by the seSTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis), insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (thrombophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention.
XX
SQ Sequence 617 BP; 165 A; 169 C; 148 G; 135 T; 0 other;
Query Match 31.3%; Score 526.4; DB 21; Length 617;
Best Local Similarity 99.8%; Pred. No. 9.5e-142;
Matches 527; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 444 GCGCGGGCCAAATGTCATTATCCACAGTCTCCCTACTACCGCTGTGCTGAACAC 503
QY 421 CCTTACACAGATGCCAGTGTCCCAAGTGTCTGTATGACAGGACCAACCCCTGCAG 480
DB 504 CCTTACACAGATGCCAGTGTCCCAAGTGTCTGTATGACAGGACCAACCCCTGCAG 563
QY 481 AATGGGGCTACTGCTCCCGCATACCGAGATCCAGTTCACCTGT 528
DB 564 AATGGGGCTACTGCTCCCGCATACCGAGATCCAGTTCACCTGT 611
RESULT 6
ABI99281
ID ABI99281 standard; cDNA; 397 BP.
AC ABI99281;
XX
XX 07-MAR-2002 (first entry)
DE Mouse ischemic condition related cDNA sequence SBO ID NO:123.
XX
XX
KW Mouse; ischemia; compressive ischemia; occlusive ischemia;
KW vasospastic ischemia; ischemic condition; ischemic disease; ss.
OS Mus musculus.
PN W0200188188-A2.
PD 22-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-JP04192.
PR 18-MAY-2000; 2000JP-0145977.
PA (UNIV-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI: 2002-034733/04.
DR
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
PS Claim 2; Page 351-352; 2690pp; English.
XX
XX The present invention describes a method for examining ischemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischemic condition (e.g. compressive
CC ischemia, occlusive ischemia or vasospastic ischemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischemic condition-improving
CC drugs or therapeutics for ischemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 397 BP; 103 A; 105 C; 111 G; 78 T; 0 other;
Query Match 14.3%; Score 241; DB 24; Length 397;
Best Local Similarity 84.4%; Pred. No. 2.6e-59;
Matches 271; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```

|||||
Db 73 CCTTTTGCACACTCCCGACACACTCTATGACCCACCATGATGACGATGATTTGTGG 132
Oy 1483 GGAATCTTTCAGAAAACCTGCGAGACACCTGCGAGGGTGACTGTGAGCCCTGACC 1542
Db 133 GGGAACTTTCAGAAACCCCGGATCGACACCTGCGAGGGTGAATCGGGGGCCCTTAAC 192
Oy 1543 TGTGAGAACGAGCGACCTACTAGCTGTATGAGATGAGTGGGGCTGGAGTGTGAG 1602
Db 193 TGTGAGAACGATGAACTTACTAGCTGTACGGATTTGTAAGCTGGGGCAGGATGTGG 252
Oy 1603 AAGAGGCGCAGGGGTCTACACCCAGTACCAATTCCTGAAATTTGATCAAGCCACATC 1662
Db 253 AAGAACCCAGGAGTCTACACCTACCAAGTCTCTGAATTTGATPAAGACCAACATG 312
Oy 1663 AAAAGTGAAGTGGCTTCTAA 1683
Db 313 CACGAGGAGCGCTGGCTTCTGA 333

```

RESULT 7
ABA58323/c
ID ABA58323 standard; DNA; 451 BP.

AC ABA58323;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #6628.

DE Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN W020015727-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human foetal liver -

XX Claim 1; SEQ ID NO 6628; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.

XX Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

XX Query Match 7.1%; Score 120.2; DB 22; Length 451;

XX Best Local Similarity 97.6%; Pred. No. 2.4e-24;

```

Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 334 TCAGAAAGTCAAAATATGTCGACAGACACCATGATGGCCGGGGCCATGTCTCATTTAC 383
Db 451 TCCTACAGTCAAAATATGTCGACAGACACCATGATGGCCGGGGCCATGTCTCATTTAC 392
Oy 384 CCAGAGTCCCTTACTACCGCTGTCTGTAAACACCTTACAGAGTCCACCTCTC 443
Db 391 CCAGAGTCCCTTACTACCGCTGTCTGTAAACACCTTACAGAGTCCACCTCTC 332
Oy 444 CCAAG 448
Db 331 CCAAG 327

```

RESULT 8
AAK06420/c
ID AAK06420 standard; DNA; 451 BP.

AC AAK06420;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 6411.

DE Human; brain expressed exon; gene expression analysis; probe;

DE microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

DE epilepsy; cancer; ss.

OS Homo sapiens.

PN W020015727-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

XX Example 4; SEQ ID NO: 6411; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

CC invention.

XX Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

XX Query Match 7.1%; Score 120.2; DB 22; Length 451;

XX Best Local Similarity 97.6%; Pred. No. 2.4e-24;

XX Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 324 TCAGAAAGTCAAAATATGTCGACAGACACCATGATGGCCGGGGCCATGTCTCATTTAC 383

Db 451 TCCTACAGTCAAAATATGTCGACAGACACCATGATGGCCGGGGCCATGTCTCATTTAC 392

OY 384 CCAGAGTCTCCCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 443
 |||
 DB 391 CCAGAGTCTCCCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 332
 OY 444 CCAAG 448
 |||||
 DB 331 CCAAG 327

RESULT 9

AAK32096/c
 ID AAK32096 standard; DNA: 451 BP.

AAK32096:

06-NOV-2001 (first entry)

Human bone marrow expressed single exon probe SEQ ID NO: 6653.

Human: bone marrow expressed exon; gene expression analysis; probe;
 microarray; cancer; leukaemia; lymphoma; myeloma; ss.

Homo sapiens.

WO200157276-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00668.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Pen SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488900/53.

Human genome-derived single exon nucleic acid probes useful for

analyzing gene expression in human bone marrow -

Example 4; SEQ ID NO: 6653; 658bp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid

probes which are derived from genomic sequences expressed in the human

bone marrow. They can be used to measure gene expression in bone marrow

samples which may enable the improved diagnosis and treatment of cancers

such as lymphoma, leukaemia and myeloma. The present sequence is one of

the probes of the invention.

Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

Query Match

Best Local Similarity 7.1%; Score 120.2; DB 22; Length 451;

Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 324 TCAGAAAGTGCAGAAATACGTGACAGACCAACCCATGTGGCGGGCCAAATGTCATTAC 383

DB 451 TCCTACAGTGCAGAAATACGTGACAGACCAACCCATGTGGCGGGCCAAATGTCATTAC 392

DB 331 CCAAG 327

RESULT 10
 AA137947/c
 ID AA137947 standard; DNA: 451 BP.
 AA137947:
 17-OCT-2001 (first entry)

Probe #6633 used to measure gene expression in human placenta sample.
 Probe: microarray; human; placenta; antenatal diagnosis;
 genetic disorder; ss.
 Homo sapiens.
 WO200157272-A2.
 09-AUG-2001.
 30-JAN-2001; 2001WO-US00663.
 04-FEB-2000; 2000US-0180312.
 26-MAY-2000; 2000US-0207456.
 30-JUN-2000; 2000US-0608408.
 03-AUG-2000; 2000US-0632366.
 21-SEP-2000; 2000US-0234687.
 27-SEP-2000; 2000US-0236359.
 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.
 Pen SG, Hanzel DK, Chen W, Rank DR;
 WPI; 2001-488897/53.

Human genome-derived single exon nucleic acid probes useful for
 analyzing gene expression in human placenta -

Claim 25; SEQ ID NO 6633; 654bp; English.

The present invention relates to single exon nucleic acid probes (SENP).
 The present sequence is one such probe. The probes are useful for
 producing a microarray for predicting, measuring and displaying gene
 expression in samples derived from human placenta. The probes are useful
 for antenatal diagnosis of human genetic disorders.

Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

Query Match

Best Local Similarity 7.1%; Score 120.2; DB 22; Length 451;

Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 324 TCAGAAAGTGCAGAAATACGTGACAGACCAACCCATGTGGCGGGCCAAATGTCATTAC 383

DB 451 TCCTACAGTGCAGAAATACGTGACAGACCAACCCATGTGGCGGGCCAAATGTCATTAC 392

OY 384 CCAGAGTCTCCCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 443

DB 391 CCAGAGTCTCCCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 332

OY 444 CCAAG 448
 |||||

RESULT 11
 ABS06863/c
 ID ABS06863 standard; DNA: 451 BP.
 ABS06863:
 ABS06863;

XX 19-AUG-2002 (first entry)
 DT
 XX
 DE Human genome-derived single exon probe from lung SEQ ID No 6854.
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsagen syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PT Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI: 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 1: SEQ ID No 6854; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes, the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsagen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp://ipo.int/pub/published_pcl_sequences.
 CC
 XX
 SQ Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;
 Query Match 7.1%; Score 120.2; DB 24; Length 451;
 Best Local Similarity 97.6%; Pred. No. 2.4e-24;
 Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 324 TCAGAAAGTCGCAAAATACGTGCAAGGACCAACCATGTGCGCGGCCCAATGTCTCATTTAC 383
 DB 451 TCCTACAGTGCAGAAATACGTGCAAGGACCAACCATGTGCGCGGCCCAATGTCTCATTTAC 392
 OY 384 CCAGAGTCTCTCCCTACTACGCGTGTGTCTTAACACCCCTTACACAGGTCCACAGTGC 443
 DB 391 CCAGAGTCTCTCCCTACTACGCGTGTGTCTTAACACCCCTTACACAGGTCCACAGTGC 332
 OY 444 CCAAG 448
 DB 331 CCAAG 327
 RESULT 12
 ABA70919/c
 ID ABA70919 standard; DNA; 117 BP.
 XX
 AC ABA70919;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #19224.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PT Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI: 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver -
 XX
 PS Claim 4: SEQ ID NO 19224; 639bp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.

CC Sequence 117 BP: 24 A; 24 C; 40 G; 29 T; 0 other;

Query Match 7.0%; Score 117; DB 22; Length 117;

Best Local Similarity 100.0%; Pred. No. 1e-23; Indels 0; Gaps 0;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 332 TGCAGAAATACGTGCAAGACACCAACCATGTGCGCGGCCCAATGTCATTACCCAGAGTC 391

117 TGCAGAAATACGTGCAAGACACCAACCATGTGCGCGGCCCAATGTCATTACCCAGAGTC 58

Qy 392 CTCCTACTACCGCTGTGTCTGTAAACACCCCTTACACAGTCCAGTGTCCCAAG 448

Db 57 CTCCTACTACCGCTGTGTCTGTAAACACCCCTTACACAGTCCAGTGTCCCAAG 1

RESULT 13

AAK19188/C
ID AAK19188 standard; DNA: 117 BP.

AC AAK19188;

DT 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe SEQ ID NO: 19179.

XX Human: brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer; ss.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX brains -

XX Example 4; SEQ ID NO: 19179; 650bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX brain. They can be used to measure gene expression in brain cell samples,

XX which may enable the diagnosis and improved treatment of nervous system

XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 TGCAGAAATACGTGCAAGACACCAACCATGTGCGCGGCCCAATGTCATTACCCAGAGTC 391

Db 117 TGCAGAAATACGTGCAAGACACCAACCATGTGCGCGGCCCAATGTCATTACCCAGAGTC 58

Qy 392 CTCCTACTACCGCTGTGTCTGTAAACACCCCTTACACAGTCCAGTGTCCCAAG 448

Db 57 CTCCTACTACCGCTGTGTCTGTAAACACCCCTTACACAGTCCAGTGTCCCAAG 1

RESULT 14

AAK45153/C
ID AAK45153 standard; DNA: 117 BP.

AC AAK45153;

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 19710.

XX Human: bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 19710; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukaemia and myeloma. The present sequence is one of

XX the probes of the invention.

XX Sequence 117 BP: 24 A; 24 C; 40 G; 29 T; 0 other;

Query Match 7.0%; Score 117; DB 22; Length 117;

Best Local Similarity 100.0%; Pred. No. 1e-23; Indels 0; Gaps 0;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 TGCAGAAATACGTGCAAGACACCAACCATGTGCGCGGCCCAATGTCATTACCCAGAGTC 391

Db 117 TGCAGAAATACGTGCAAGACACCAACCATGTGCGCGGCCCAATGTCATTACCCAGAGTC 58

Qy 392 CTCCTACTACCGCTGTGTCTGTAAACACCCCTTACACAGTCCAGTGTCCCAAG 448

Db 57 CTCCTACTACCGCTGTGTCTGTAAACACCCCTTACACAGTCCAGTGTCCCAAG 1

```
RESULT 15
AAI51106/C
ID AAI51106 standard; DNA: 117 bp.
XX
XX AAI51106;
AC
XX
XX 17-OCT-2001 (first entry)
DT
XX
DE Probe #19792 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200157272-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00663.
PE
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-488897/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta.
PT
XX
XX Claim 25; SEQ ID No 19792; 654bp; English.
PS
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
XX
XX Sequence 117 bp; 24 A; 24 C; 40 G; 29 T; 0 other;
SQ
Query Match 7.0%; Score 117; DB 22; Length 117;
Best Local Similarity 100.0%; Pred. No. 1e-23;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 332 TGCATAATACGTGCAAGACACCATGTGGCGGGCCCAATGTCATATTACCCAGAGTC 391
DB 117 TGCATAATACGTGCAAGACACCATGTGGCGGGCCCAATGTCATATTACCCAGAGTC 58
OY 392 CTCCTACTACCGCTGTCTGTAAACAACCTTACACAGGTCACAGCTCTCCCAAG 448
DB 57 CTCCTACTACCGCTGTCTGTAAACAACCTTACACAGGTCACAGCTCTCCCAAG 1
Search completed: March 6, 2003, 16:01:40
Job time : 348.5 secs
```